Genetic, physiological and ecological diversity of the diatom genus *Leptocylindrus*

Thesis

The Open University, London
Stazione Zoologica Anton Dohrn,
Naples

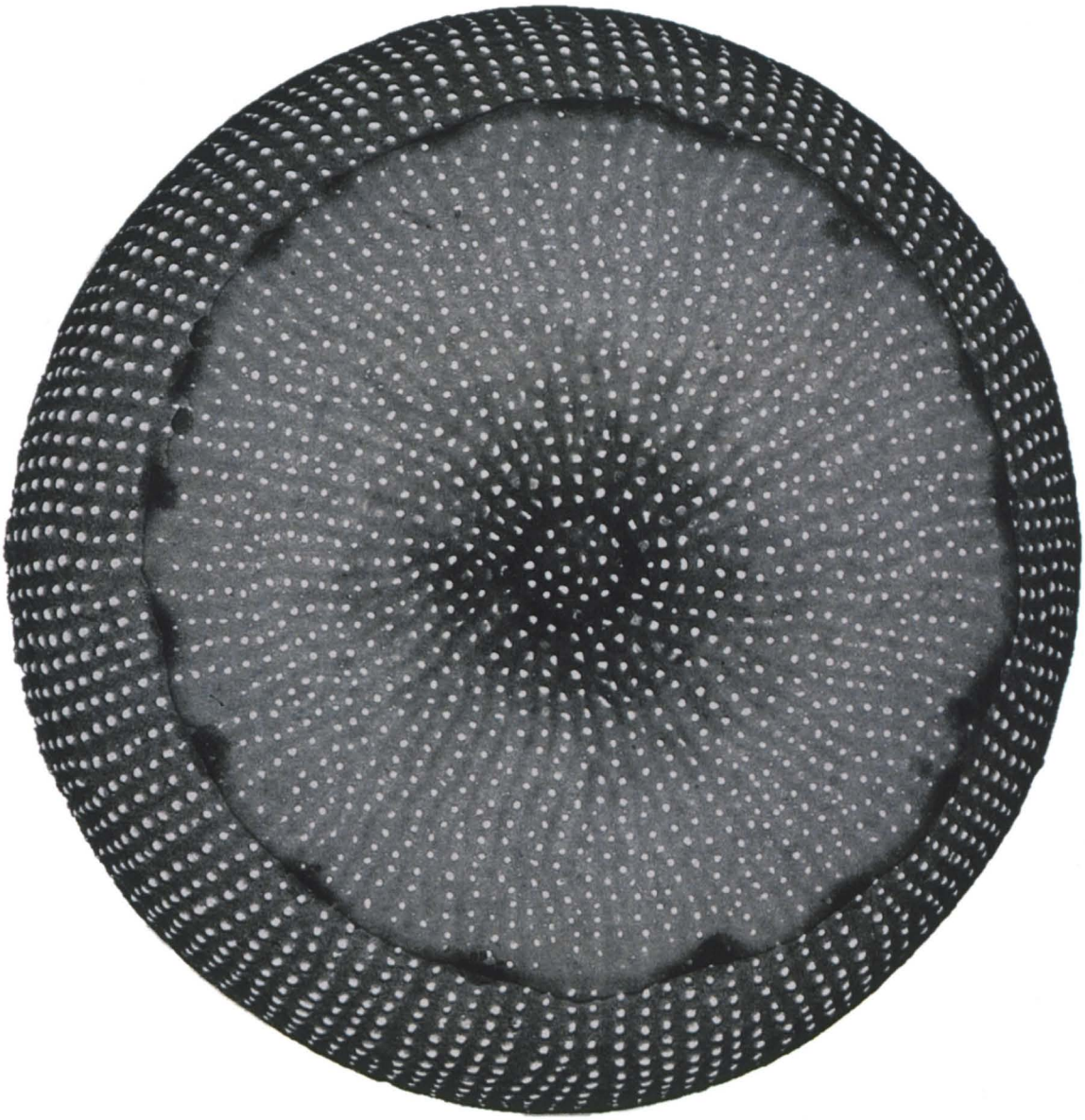
Genetic, physiological and ecological diversity of the diatom genus *Leptocylindrus*

DEEPAK NANJAPPA

Thesis submitted for the degree of Doctor of Philosophy (PhD) in Life and Biomolecular Sciences

September 2012

DATE of SUBMISSION : 30 SEPTEMBER 2012
DATE OF AWARD: 21 DECEMBER 2012



Dedicated to my beloved parents and family

Take up one idea. Make that one idea your life - think of it, dream of it, live on that idea. Let the brain, muscles, nerves, every part of your body, be full of that idea, and just leave every other idea alone. This is the way to success. – Swami Vivekananda

Thesis supervision committee

Director of Studies

Dr Adriana Zingone
Senior Scientist
Stazione Zoologica Anton Dohrn
Villa Comunale, Naples, Italy

Internal Supervisor

Dr Wiebe H.C.F. Kooistra
Senior Scientist
Stazione Zoologica Anton Dohrn
Villa Comunale, Naples, Italy

External Supervisor

Professor David G. Mann
Senior Principal Research Scientist
Royal Botanic Garden Edinburgh
Edinburgh EH3 5LR
Scotland, UK

Table of contents

YOUR ACCEPTANCE

1 Student details

Your full name: Deepak Nanjappa

Personal identifier (PI): B2978976

Affiliated Research Centre (ARC) (if applicable): Stazione Zoologica A. Dohrn

Department: Ecology and Evolution of Planktons

Thesis title: Genetic, physiological and ecological diversity in the diatom genus Leptocylindrus

2 Authorisation statement

I confirm that I am willing for my thesis to be made available to readers by The Open University Library, and that it may be photocopied, subject to the discretion of the Librarian

Signed: .....

Print name: Deepak Nanjappa

Date: 21/01/2013 DD/MM/YY

3 British Library Authorisation (PhD and EdD candidates only)

The Open University has agreed that a copy of your thesis can be made available on loan to the British Library Thesis Service on a voluntary basis. The British Library may make the thesis available online. Please indicate your preference below:

☒ I am willing for The Open University to loan the British Library a copy of my thesis

OR

☐ I do not wish The Open University to loan the British Library a copy of my thesis

Abstract

Diatoms are among the most productive photoautotrophic organisms on Earth. Within the diatom genus *Leptocylindrus*, two species, *L. danicus* and *L. minimus*, are reported as abundant in coastal waters worldwide and in the Gulf of Naples (GoN). This thesis aimed at a closer characterisation of *Leptocylindrus* species through different approaches, including the study of their morphology, molecular phylogeny, metagenomics and biochemistry. Eighty-three strains from the GoN and one strain (CCMP 1856) from the Atlantic US coast were analysed. Based on the morphological, molecular and life cycle differences, the taxonomy of the genus was revised. The GoN species previously identified as *L. minimus* was in fact identified as *L. belgicus* Meunier and was placed in a new genus, *Tenuicylindrus* Nanjappa and Zingone which, along with *Leptocylindrus* Cleve, belongs to the family Leptocylindraceae. Five species in all were included in the genus *Leptocylindrus*: *L. danicus* Cleve and *L. minimus* Gran, two novel species *L. hargravesii* and *L. convexus*, and *L. aporus* (Hargraves) Nanjappa & Zingone, which was raised from the variety to the species status. The real *Leptocylindrus minimus* was not found in the GoN. To address the distribution of the 6 species outside the GoN, two metagenomic databases, BioMarKs (Europe) and *Tara* Oceans (worldwide) were explored. Sequences of *L. aporus*, *L. convexus* and *L. danicus* were recovered at many sites in European waters and across the world's seas, while those of *L. minimus* were retrieved only in the Oslo fjord and those of *T. belgicus* were only found in the GoN and Oslo fjord. Additional diversity was observed in the *Tara* Oceans dataset but, in lack of morphological information, whether this diversity is real remains to be clarified. All species except *L. minimus* were also categorised based on the diversity in their oxylipins pathways. *Leptocylindrus danicus* and *L. hargravesii* shared common lipoxygenase pathways, different from the ones shared by the species *L. aporus* and *L. convexus*. *Tenuicylindrus belgicus* exhibits a pathway distinct from that of *Leptocylindrus* species. Species-specific compounds produced in minor quantities were also observed. Physiological experiments show that *L. aporus* can withstand higher (26 °C) temperature but not lower (12 °C) temperature, while *L. danicus* can withstand low temperature but not high temperature. This corresponds to the species occurrence in the natural environment, where *L. aporus* blooms during summer and *L. danicus* is found in all seasons except summer. Altogether, through an interdisciplinary approach, the studies described in this thesis provide substantial information that may have important implications in the field of ecology, evolution, conservation biology and biotechnology.

Acknowledgments

During the three years of my PhD, I have been inspired, advised and supported by a number of people who have contributed in the successful completion. I wish to express my sincere gratitude to all of them.

It gives me great pleasure to express my sincere gratitude to Dr Adriana Zingone, director of studies, for giving me the opportunity to be her student, inspiring me towards the ecological and evolutionary based concepts and for the guidance that I received throughout the period. I am grateful to her for giving me freedom in designing the project, critically evaluating my ideas and proposals. These have helped me to learn to formulate collaborative projects with people sharing diverse interests.

I am also grateful to my internal supervisor Dr Wiebe H.C.F. Kooistra, who not only has guided me in all the phylogenetics research undertaken but has introduced and inspired me to many other evolutionary concepts and geological time scale of events. Thanks Wiebe for also giving me many of the issues of The Economist, Scientific American and National Geographic journals that helped me to keep track of the latest activities. I can never forget the terrace parties with different cuisine, speaking about planetary positions, politics, religious beliefs and many more interesting topics that made me to ponder on those issues.

Sincere thanks to my external supervisor, Professor David G. Mann, for the intense and fruitful discussions, his questions and suggestions have helped me to make important decisions. I was inspired by enthusiasm, simplicity, kindness and thought process. I consider it an honour to have him as my PhD supervisor. I am grateful to him!

I would like to thank also our laboratory coordinator Dr Marina Montresor from whom I received many suggestions and advices. Her ideas, time management skills and dedication to research keep inspiring me. Many thanks to all the staff and colleagues people in LEEP especially to Dr Diana Sarno, Dr Christophe Brunet and Dr Mariella Ferrante. Thanks to Dr Maria Valeria Ruggerio who not only gave advice but also rented her apartment at an economic price.

I am also grateful to two of the collaborative groups that I collaborated for the metagenomics and metabolomics projects undertaken during the study. Dr Chris Bowler and his team at Ecole Normale Supérieure, Paris for metagenomic study. Many thanks to Dr Stephane Audic at Station Biologique de Roscoff, who was also part of metagenomic project I under took

during my PhD. Dr Angelo Fontana and Dr Guliana d'Ippolito at *CNR-Instituto di Biochemica e Biomoleculare*, Naples for their support in biochemistry work.

Thanks to Gandi Forlani for teaching me electron microscopy, Alessandro Manfredonia for the support in daily laboratory activities and Carmen Minucci for her technical help.

It is never complete without acknowledging friends, for the fun, aperitif and parties all memorable events. Thanks to all my friends Eleonora, Laura, Sylvie, Guiseppe, Alessandra, Cecilia, Vasco, Agostina, Davide, Anna, Cristina, Swaraj, Krzystof, Gauri, Ferdinando, Luciana, Gianluca, Filip, Ankita, Atul, Praveen, Raghu, Sneha, Shrikant, Ashwani, Anupum and many more friends those I have met, shared thoughts and had fun. They are all important part of my PhD life in Naples.

I also would like to thank all the people that I met during this period: my sincere thanks for sharing with me your free time! I am glad to have the opportunity to live in the beautiful, historic and lively city, Naples with lots of fun. Its home to the best pizzas in the world, thanks for giving such tasty pizzas.

Special thanks to the Stazione Zoologica Anton Dohrn for accepting me as PhD student, providing financial support and for giving remarkable PhD life in its diverse learning environment.

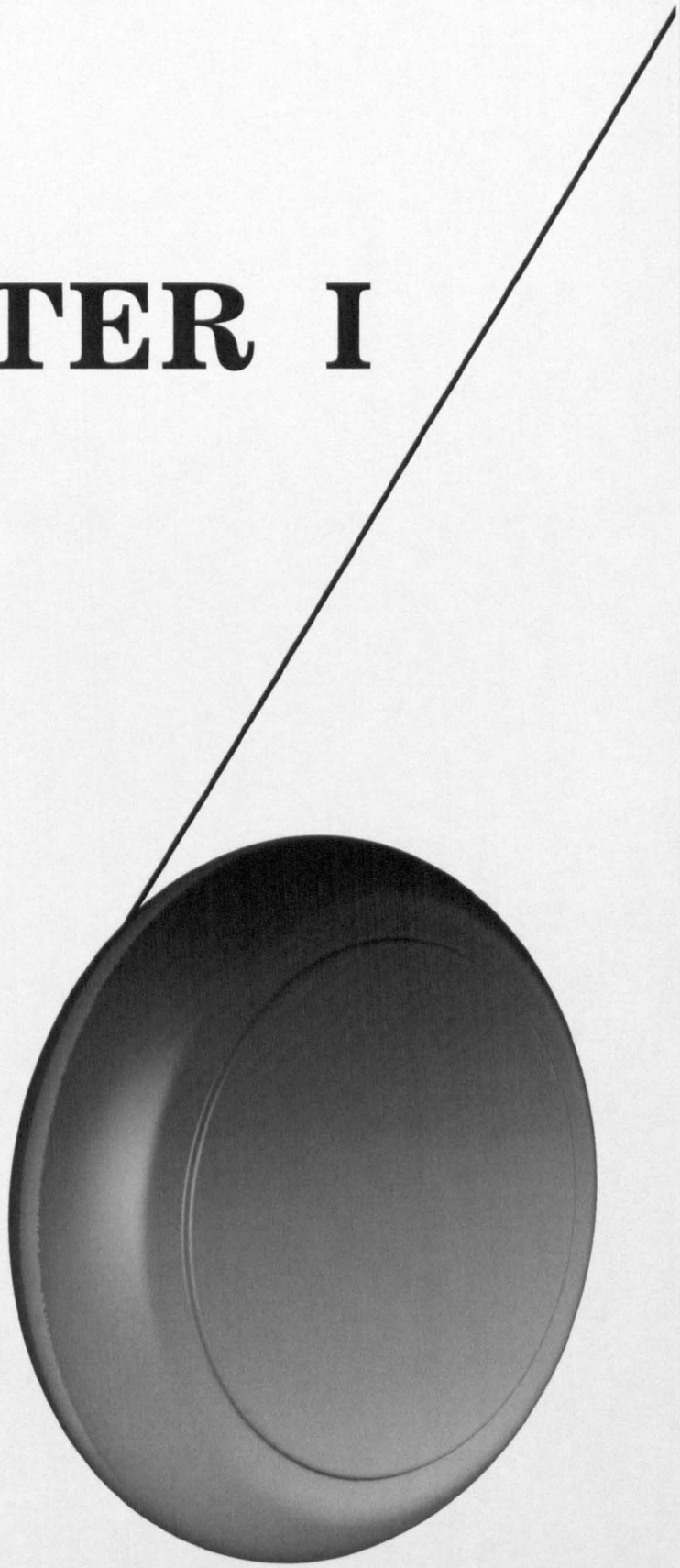
List of tables

List of figures

Abbrevations

μ_{max}	Maximum Growth Rate
AIC	Akaike Information Criterion
BioMarKs	Biodiversity of Marine Eukaryotes
BLAST	Basic Local Alignment Search Tool
CBC	compensatory base changes
C_{max}	maximum cell density
COX	Cytochrome c oxidase
DCM	Deep Chlorophyll Maximum
DHA	Docosa Hexaenoic Acid
DMS	Dimethylsulfide
DMSP	Dimethylsulfoniopropionate
DOC	Dissolved Organic Carbon
EPA	Eicosa Pentaenoic Acid
FAH	Fatty Acid Hydroperoxides
FISH	Fluorescent In Situ Hybridization
GCMS	Gas Chromatography Mass Spectrometry
GoN	Gulf of Naples
LCMS	Liquid Chromatography Mass Spectrometry
LM	Light Microscope
LOX	Lipoxygenase
LTER-MC	Long term ecological research - MareChiara
MES	Mesopelagic
ML	Maximum Likelihood
MP	Maximum Parsimony
NCBI	National Center for Biotechnology Information
NCMA	National Center for Marine Algae and Microbiota
NJ	Neighbour Joining
NMR	Nuclear Magnetic Resonance Spectroscopy
OMZ	Oxygen Minimum Zone
OTU	Operational Taxonomic Unit
PAUP	Phylogenetic Analysis Using Parsimony
psbC	photosystem II chlorophyll a-binding protein coding gene
PUA	Poly Unsaturated Aldehydes
PUFA	Poly Unsaturated Fatty Acid
rbcL	Ribulose-Bisphosphate Carboxylase gene
SEM	Scanning Electron Microscope
SFA	Saturated Fatty Acids
TEM	Transmission Electron Microscope
USFA	Unsaturated Fatty Acids

CHAPTER I



General Introduction

1.1 Phytoplankton

The Ocean covers 71 percent of the Earth's surface and contains 97 percent of the planet's water in the global hydrological cycle. It is home to three quarters of all life on the Earth, mostly unexplored. Among the great diversity of life the sea supports, there is a quantitatively and ecologically extremely important group of microscopic organism that, all together, are responsible for most of the marine primary production. Taken together, these organisms constitute the marine 'Phytoplankton'. The term Phytoplankton comes from the Greek words, *phyton,* which means "plant" and plankton, which means "wanderer" or "drifter."

The marine phytoplankton is responsible for the bulk of the primary production in the sea, generating biomass from simple inorganic molecules using light energy. Macroalgae and seagrasses are, of course, also relevant primary producers in the marine realm, but these organisms are confined essentially to the photic zone of coastlines and as drifters in the water surface. Instead, the phytoplankton occupies the photic zone of the entire ocean. Through its photosynthetic activity, the phytoplankton contributes to half the oxygen present in the earth's atmosphere and are responsible for approximately half of the global (terrestrial and marine) net primary production (Field *et al.* 1998), (Nelson *et al.* 1995). In addition to their ecological relevance, microalgae have potential biotechnological applications, e.g., as potential producers of biofuel and as systems for heterologous protein expression. These photosynthetic 'workhorses' can convert solar energy more efficiently (5% as compared to 1.5 – 2% in land plants) and have much higher lipid yields (may exceed 70%, w/w) than those of agricultural oleaginosous crops (5% w/w).

Under suitable environmental conditions microalgae can experience elevated growth rates and attain high cell densities. Such bursts of phytoplankton growth are usually

composed of many different species and are commonly referred to as algal blooms (Diersing 2009). Phytoplankton blooms are governed by nutrient availability and external physical factors such as temperature and light. For example during spring at temperate latitudes, there is a burst of microalgal growth in response to increasing light availability and water column stabilisation (Sverdrup 1953), which is defined as the spring bloom, the spring bloom usually declines because nutrients are depleted in the photic zone and the thermocline prevents mixing of nutrient-rich deep water. Another bloom occurs in the temperate zone during the autumn because the breakup of the thermocline allows a replenishment of nutrients in the photic zone whilst there is still enough light to permit net growth in the mixed water layer. In the monsoon-governed parts of the oceans phytoplankton blooms are governed by the nutrient-rich runoff during the wet season and due to upwelling of deep nutrient-rich water due to trade winds. These blooms are natural events in the annual phytoplankton cycle of these various regions.

Phytoplankton species reacts quickly to environmental fluctuations (e.g. nutrient concentration, temperature values and light irradiance) because of small cell size and high growth rates. This rapid response allows ecological studies on the reaction of the various species or of the entire community to physical-chemical processes occurring at a wide range of spatial and temporal scales (Harris 1986). This task is obviously easier if phytoplankton is considered as a unique entity, but becomes much more difficult to tackle if individual species responses to environmental fluctuations are addressed.

Phytoplankton comprises a diverse group, incorporating protistan eukaryotes as well as eubacterial and archaebacterial prokaryotes. The most important groups of phytoplankton, in terms of biomass, includes the diatoms, the cyanobacteria, the haptophytes (including the coccolithophores) and the dinoflagellates, although at times, other flagellate groups can contribute significantly to phytoplankton blooms as well.

Among the other eukaryotic phytoplankton are the chlorophytes, cryptomonads, chloromonads, chrysomonads and euglenoids.

Cyanobacteria are the most important representative among all phototrophic forms of bacteria and include four main evolutionary lines (the chroococcalean, the oscillatorialean, and the nostocalean; the stigonematalean line). While prokaryotic phytoplankton such as *Procholorococcus* are the most widespread (Moore *et al.* 1998) the greater part of species diversity is found in the diatoms group.

1.2 Diatoms

Diatoms are among the most common types of phytoplankton. Diatoms constitute the class Bacillariophyceae within the photoautotroph Stramenopiles. The name diatom is derived from the Greek word, *diatomos* meaning 'split into two'; it refers to their distinctive two-part cell walls that fit each other. Diatoms are strictly unicellular algae characterized by unique compound-cell walls of amorphous silica (early diagrams by Haeckel 1873, Fig 1.1). Many species form chains composed of sister cells, but the cells in such chains do not retain cytoplasmatic contact with one another.

Fossil evidence suggests that diatoms originated during the Jurassic period (Kooistra and Medlin 1996, Schieber *et al.* 2000). Since then they have been found in fossil strata of marine and freshwater habitats. Diatoms can now be found in large numbers in almost all waters of the world, both freshwater and saline, as well as in temporarily moist soil and on wet surfaces (Kooistra *et al.* 2007). Although many species occur in the plankton, the majority of the diversity abounds attached to or gliding over surfaces, or drifting over sediments. A few species occur endosymbiotically in dinoflagellates and microzooplankton. Under favourable conditions, planktonic diatoms "bloom", reproducing clonally rapidly so that they dominate phytoplankton populations. When conditions turn less favourable, for instance when light levels become too low, or when

one or more critical nutrients get depleted, they can turn dormant and sink to lower levels until currents return them to the surface and a new bloom can take place.

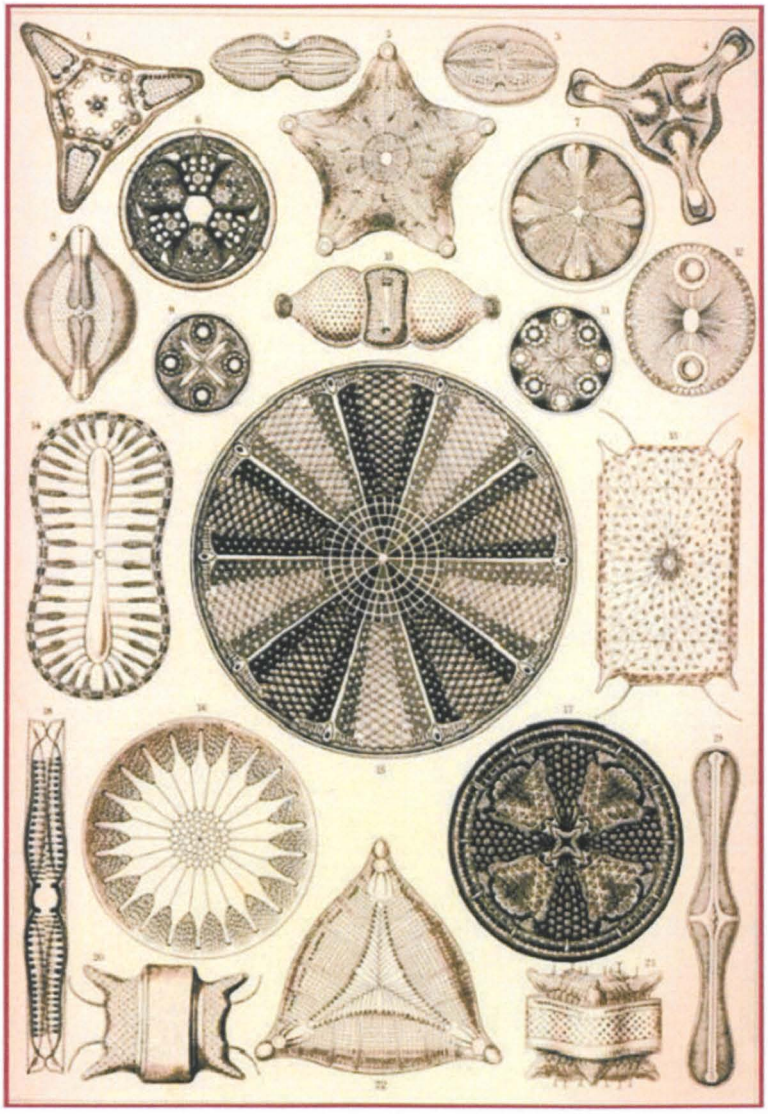


Fig. 1.1. Early drawings of phytoplankton diatoms (Haeckel 1873).

Diatoms constitute approximately 40% of the phytoplankton and play a major role in the ecology and biogeochemistry of the Earth. Diatoms alone contribute approximately 40% of the 45–50 billion metric tonnes of organic matter generated annually in the ocean by photosynthesis (Nelson *et al.* 1995). Thus constitute a key component of the biological carbon pump that transports carbon to the seafloor, where it can get sequestered in the sediment and hence, fossilized. In this way diatoms contribute to long-term CO_2 sequestration.

1.3 General biology

A characteristic hallmark of diatom cells is that they are encased within a highly nanopatterned, composite cell wall known as a frustule. The frustule elements are made of amorphous silica (hydrated silicon dioxide, $[(SiO_2)_n(H_2O)]$), laid down upon an organic matrix (Round *et al.* 1990). Beneath the frustule is the protoplast, which consists of a large vacuole, a diploid nucleus and one to several plastids. In its simplest form, a cell wall is composed of two valves, the epivalve and the hypovalve, accompanied by series of girdle bands (also called the cingular bands) (Round *et al.* 1990). Each valve in its simplest form looks like a petri dish with a flat area, called valve face, and a rim connecting the girdle, called mantle. The first girdle band that connects the valve is called the valvacopula, the other bands are referred to as copulae (also, intercalary bands or pleurae). The valve is often ornamented with pores (areolae), processes, spines, hyaline areas and other distinguishing protrusions. They might provide an added advantage to the cell either protection from grazers or increasing the buoyancy. These patternings are so precise and well maintained over mitotic cell divisions that, they have attracted materials scientists interested in silica-based nanoparticles (Gordon *et al.* 2009). The frustule is coated with a layer of organic substance, sometimes pectin. Heterokont algae have chloroplasts that are surrounded by four membranes (Gibbs 1979). Counted from the outermost to the innermost membrane, the first membrane is continuous with the host's chloroplast endoplasmic reticulum, or cER. The second membrane presents a barrier between the lumen of the endoplasmic reticulum and the chloroplast, which represents the next two membranes, within which the thylakoid membranes are found. This arrangement of membranes suggests that the diatom plastid results from two endosymbiosis events. In the first of such an event, a eukaryotic heterotroph engulfed a cyanobacterium, giving rise to all of the eukaryote autotrophs. The innermost membrane is derived from this cyanobacterium (Reyes-Prieto

et al. 2006). In a secondary event, a eukaryotic heterotroph incorporated a red alga as endosymbiont. The second of the four membranes is derived from the plasmalemma of this primary host and the third of the plasmalemma of the red alga, the outermost membrane is then the plasmalemma of the secondary host. This arrangement gave rise to a lineage including the Stramenopiles and probably also the haptophytes, the cryptophytes, the picobiliphytes and the dinoflagellates. Stramenopiles include diatoms, a series of brown and golden-brown microalgae, the brown macroalgae as well as parasitic and saprotrophic marine fungi and parasites of land plants (Gibbs 1981, Parker *et al.* 2008, Cavalier-Smith 1993, Cavalier-Smith 2000). More recently, genomic data have suggested a more complex evolutionary history for diatoms involving a series of endosymbiotic events some of which totally unexpected (Sanchez-Puerta & Delwiche 2008). A first surprise was the discovery of a bacterial (chlamydial) endosymbiosis prior to the primary endosymbiotic event. A subsequent genome-wide gene derivation study revealed that many genes (>1700, constituting 16% of diatom nuclear coding potential) show high similarity with those from green algae, suggesting an additional green algal endosymbiotic event (Moustafa *et al.* 2009). Finally, bacteria and viruses also seem to have mediated gene transfer adding further complexity to the genome of diatoms (Montsant *et al.* 2007).

The chloroplasts characteristically contain chlorophyll a and chlorophyll c, and usually the accessory pigment fucoxanthin, giving them a golden-brown or brownish-green colour. "Diatoms exhibit greater variation in chloroplast morphology and arrangement than any other group of algae...." (Mann 1996a).

1.4 Diatom classification

In traditional publications, diatoms have been classified taxonomically into two main classes according to the shape and symmetry of their cell walls; this includes centric diatoms or centrales and pennate diatoms or pennales.

The more recent classification divides diatoms into three classes (Round *et al.* 1990):

1.4.1 Centric diatoms (Coscinodiscophyceae). Centric diatoms are either radially symmetric or bi/multipolar whereas pennates are asymmetrical. Radial centrics have a petri-dish like architecture with a circular centre of symmetry (annulus) and rows of pores (striae) radiating from the annulus (Fig. 1.2, A). From the phylogenetic viewpoint they are considered the most ancient group. The multipolar centric diatoms constitute a newer group, which emerged from radial centric ancestry (Kooistra *et al.* 2007). Polar centrics have an elongated or distorted annulus (Fig. 1.2, B) while no raphe or pseudoraphe is present and spontaneous movement does not occur. From this group the pennates evolved.

Centric are normally planktonic, and several species (e.g in the genera *Chaetoceros*, *Thalassiosira*, *Skeletonema* and *Leptocylindrus* often constitute a major part of the phytoplankton bloom in coastal ecosystems and upwelling regions.

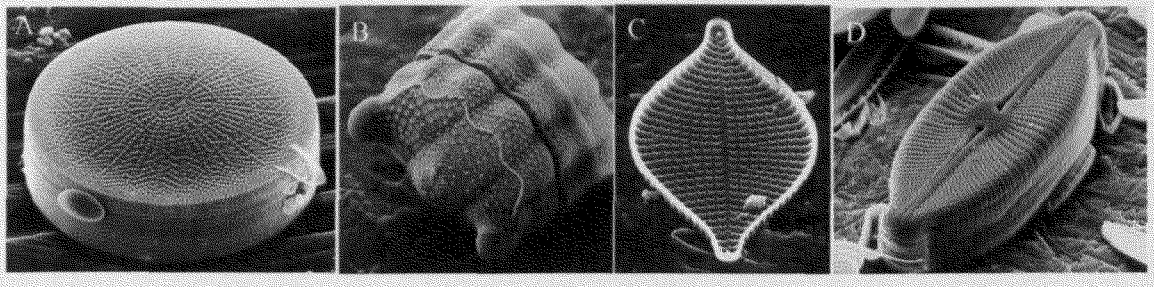


Fig. 1.2. Scanning electron microscopy images of the cell walls of four different diatom species (taken from 'Tree of life' web; copyright David G. Mann). (A) *Actinocyclus;* radial centric diatom. (B) *Biddulphia;* polar centric diatom. (C) *Rhaphoneis amphiceros;* araphid pennate diatom. (D) *Cosmioneis pusilla;* raphid pennate diatom.

1.4.2 Pennate diatoms. Pennates are bilaterally symmetrical, and their annulus is replaced by a rib (sternum), running along the longitudinal axis of each valve. Based on the presence or absence of a slit (raphe), pennates are classified into raphid or araphid, respectively.

- without a raphe (Fragilariophyceae; Fig. 1.2, C)

- with a raphe (Bacillariophyceae; Fig. 1.2, D)

Valves of both centrics and pennates may possess additional structures depending on the taxonomic group, for example most of the centrics and pennates possess rimoportulae i.e., processes with a lip-like (labiate) extension on the interior side. Pennates are often benthic, living on sediments or as epiphytes on macrophytes or invertebrates. However, some pennate diatoms are successful in the plankton, where they often persist in the open ocean waters as well as in turbulent coastal waters, demonstrating their ability to tolerate a wide range of environmental conditions. Among pennates those belonging to the genus *Fragilariopsis,* specifically *F. kerguelensis,* are prominent and form an important component of polar seas and especially the HNLC (High Nutrient Low Chlorophyll) of the Southern Ocean despite the iron limitation of the area. These diatoms escape grazing by building thick-walled frustules (Zielinski and Gersonde1997) and thus outcompete other diatoms contributing substantially to the diatom ooze accumulating under this region. Their local dominance has consequences for the capability of the ecosystem to sequester CO_2 (Treguer *et al.* 1995).

1.5 Life cycle features

Diatoms have a diplontic life cycle (diploid cells undergo growth), unique among algae (Al-Kubaisi 1981) where the vegetative cells are diploid (2n) and their gametes haploid (n) (Round *et al.* 1990, Chepurnov *et al.* 2004). By far the dominant part of the life cycle, diatoms grow vegetatively through mitosis followed by cytokinesis. Vegetative growth is usually very fast in diatoms with a concomitant increase in biomass, generally showing higher growth rates than other algae of comparable size (Furnas 1990). This is possibly due to their silica cell wall because it has been shown that the synthesis of the frustule requires less energy per atom incorporated compared to the synthesis of organic cell walls (around 8%) (Raven 1983, Raven and Waite 2004)).

Cell division starts with chloroplast division. Chloroplasts divide either autonomously, in which case chloroplast constriction occurs independent of other organelles, or their division is imposed as a consequence of cell division (cytokinesis) (Mann 1996b). After cytokinesis, but before the daughter cells separate, each of the daughters acquires one parental theca and synthesizes a new valve and, in some of the species, an accompanying set of girdle bands. Valves are formed only during cell division and the cell volume can increase during interphase by the stepwise synthesis of girdle bands. The new valve — the hypovalve of the new cell — develop within the confines of the parental theca. As a consequence, the sibling cell that inherits the parental hypotheca is usually smaller than the parental cell, and consequently, the average cell size decreases and the variance in cell sizes increases with continued vegetative growth (the MacDonald-Pfitzer rule (Round *et al.* 1990a).

Populations escape from this miniaturization trap through sexual reproduction. At a critical size threshold, they become sexually mature and can engage in sexual reproduction. The capacity of cells to become sexualized has been demonstrated to be size dependent. However, being in the right cell size window is necessary but often not sufficient condition for the induction of the sexual phase and environmental factors and/or chemical interactions between cells can further regulate the process (Chepurnov *et al.* 2004). Gametogenesis occurs by meiotic cell divisions and the cell walls of the gametangia (gamete-producing cells) are discarded. The resulting gametes fuse to form a zygote, which expands twice or three times the original cell volume, mainly through vacuolar expansion, to form a specialized cell, called auxospore (Chepurnov *et al.* 2004). The auxospore develops an organic wall, or more commonly, an organic wall that has siliceous components embedded within it or beneath it. The types of component depend on the group to which the species belongs. Multipolar centric diatoms and pennate diatoms form sets of bands called properizonial or perizonial bands, which constrain expansion of the auxospore into a non-radial shape. As a result, the diatom

emerging from this cell is multipolar or elongated as well. Cells that fail to undergo sexual reproduction and cell size restitution continue to divide mitotically until they become critically small and eventually die.

An alternate possibility to restore cell size, that has been observed in some diatom species, namely through a non-sexual process called vegetative enlargement. This phenomenon implies the discard of both valves and synthesis of new bigger ones. Vegetative enlargement has been observed in some centric diatoms (Gallagher 1983) whereas in pennates it is exceptional and appears to occur only under extreme conditions (Chepurnov and Mann 1997).

1.6 The study model *Leptocylindrus danicus*

Leptocylindrus danicus Cleve (1889) is a centric diatom belonging to the class Coscinodiscophyceae, order: Leptocylindrales, family: Leptocylindraceae. The species is widespread in coastal ecosystems and has been reported frequently along continental shelves where it constitutes one of the dominant blooms forming species. (Round *et al.* 1990) describes the genus *Leptocylindrus* as follows:

> "Cell narrow, cylindrical, united to form filaments. Chloroplast 2- many, plate like, lying against the girdle. A common, delicate member of the marine plankton with less than five species. Valves are circular; thin with a ring of projections around the margin of the flat valve face, beneath this is a shallow valve mantle bearing vertical rows of areolae. The marginal projections appear flap-like between these are other blister –like markings. Valve face with indistinct areolae in uniserate striae, these radiate from a central annulus within which is a central cluster of irregularly placed areolae. There is also an off-centre pore whose nature is uncertain and which has no parallel in any other genus. The flaps on the edge of the valve face have no structural counterpart on the inside of the valve, but in this region the radiating 'ribs' of

the valve face become more distinct before running down the valve mantel; this feature is especially obvious in TEM. Copulae many, consisting of collar-like segments, each finely porous and with a wide truncated 'ligula'".

The understanding of the taxonomic history of *L. danicus* starts with a misinterpretation of the frustule structure by Cleve (1889) during his expedition at Kattegat Bay (Fig 1.3). The original description says the species

"has convex shape, with pretty solid valves, as the circumference is nearly circular and does not show any punctuations or striations. The conjunctiva is very thin and lacks the characteristic of Rhizosolenia annular condensations. Cells were found in long chains."

Fig. 1.3. Redrawn pencil diagram of *Leptocylindrus danicus* Cleve (1889). Sketch of three colonial cells. Note the shape of valve face.

In his drawings he shows the valves as convex, which is rather unusual for *Leptocylindrus*. Later, Cleve (1894) modifies the original description of the species as it was, in his words,

"Originally described from burnt and somewhat misshaped specimens".

This error apparently mended and re-description is as follows (Fig. 1.4);

"Cells cylindrical, with flat (dried convex) ends, forming filaments. Valves without processes or perceptible structure. – Connecting zone very thin, without annuli – Cell contents: a few scattered granular chromatophores. Diameter of the filament 0.01 mm. Length of the cell 0.03 to 0.06 mm".

Fig. 1.4. Original pencil diagram of mended *Leptocylindrus danicus* Cleve (1894). Sketch of four cell colony.

Gran (1912) undertook a study in the Kattegat bay and described *Leptocylindrus* resting spores as spiny and semi-circular. Adding to the species diversity, French and Hargraves (1986) described a novel variety of *L. danicus* based on subtle morphological difference and evident life cycle variations. He found the cells sufficiently similar so as to be considered as a variety of the species *L. danicus* and named it as *L. danicus var. apora* (French III and Hargraves 1986). The new variety differed from the original *L. danicus* in that absence of sub central pore and followed vegetative cell size restoration instead of the sexual cycle observed in the nominal species.

In 1915, Gran was the first to distinguish and identify a second species in the genus *Leptocylindrus.* He described *L. minimus* (Fig. 1.5) based on water-mounted sample as:

"Chains of cylindrical cells 25μ long and 1.5–2.5 μ in diameter, each cell containing two chromatophores, one on either side of the cell nucleus. The chains are straight or slightly twisted".

Fig. 1.5. Original pencil diagram of *Leptocylindrus minimus* Gran H.H. (1915). Sketch of four cell colony.

The important distinguishing characteristic of species is the presence of two elongate chloroplasts that accommodate in a narrow cell of 2μm. Considering the only difference in the size of cell diameter and the possibility of having narrow *L. danicus,* Riley and Conover (1967) expressed the doubt as to whether *L. minimus* has to be considered as a separate species. Hasle (1975) also expressed doubts to consider *L. minimus* as separate species since the valve structure analysed by TEM was very similar to *L. danicus*. This was clarified by further studies on the species, which defined the morphology and distribution in comparison to *L. danicus* (Hargraves 1990). *Leptocylindrus minimus* has

the following key features, valve diameter of 1–4 µm and tightly bound to its adjacent band of the cingulum. The valve face is ornamented with radial rows of simple poroid areolae radiating from the central area. The central area is normally thicker than those of *L. danicus*. Along the valve margin a circumferential ring of raised blunt spines are present whose number varies from 3–5 µmin 1 µm. The bands of the cingulum are described to be half and are similar to that of valvacopula. The valvacopula junctions exhibit finely serrate margins. A concluding evidence of the species identity was provided by the discovery of the resting spore morphology. The distinguishing features in the two species were that *L. minimus* has always two chloroplasts and sometimes one when the valve diameter approaches 1 µm. Whereas *L. danicus* always has multiple chloroplasts and very rarely could have two in *invitro* cultures, however this overlie was not recorded in natural samples. On the whole, the studies have established that life cycles of the two species with similar morphology and resting spore structure and formation.

In the same year in which *L. minimus* was described, Meunier (1915), described a new species *L. belgicus* from the Flemish water, a species with similar morphology of *L. minimus.* The description states,

> "extremely narrow frustules (2 µ wide in average), arranged in a colony straight, rigid, free from differentiation noticeable. Chromatophores few, elongated in the longitudinal direction frustules."

This description is lot similar to *L. minimus* hence it was later considered to be the synonym.

Leptocylindrus mediterraneus is another species appearing in phytoplankton identification guides. The species was originally described as *Dactyliosolen mediterraneus,* is studied a lot because of its unique morphological and biological characteristics. The species is mostly associated with a protozoan epiphytic, *Solenicola setigera.* Based on these unique characteristics, the identity of this species as a diatom

has at times been questioned. Other species generally attributed to the genus include *L. adriaticus* Schröder and *L. curvatus* Skvortzov. These descriptions are imprecise, rendering identification of specimens challenging.

1.6.1 ***The life cycle of*** **Leptocylindrus.** According to the McDonald-Pfitzer rule in most diatoms valve diameter decreases from generation to generation. After they reach the minimal size, original size is restored either through sexual reproduction or through vegetative expansion. Sexuality and other size regenerative processes are very important stages of the life cycle of diatoms. The knowledge of life cycle of the diatom *L. danicus* comes mainly from the studies of French and Hargraves (1985, 1986) and Davis *et.al* (1980). These studies have contributed to the understanding of unique sexual life cycle specific to species (Fig. 1.6.1, a–m) and vegetative cell size expansion process (Fig. 1.6.2, a–i). The formation of resting spores in phytoplanktons is believed to be a strategy to survive under adverse environmental conditions. Centric diatoms are well known for the formation of resting spores both in the environment and in *invitro* cultures. Formation of highly silicified resting spores in *L. danicus* was first reported by Murray and Hjort (1912), whereas Gran (1915) reported that they develop directly from auxospores. Subsequently, Davis et al. (1980) and French & Hargraves (1985, 1986) demonstrated the formation of resting spores in *in vitro* cultures. These studies have provided a great deal of data to understand the life cycle and spore formation in *L. danicus*: 1) the environmental conditions which induce resting spore formation, and 2) those that foster their germination, 3) the physiological processes that are involved in the formation and germination of spores and, 4) capabilities of spore to withstand adverse climatic conditions. Resting spore formation has been attributed mainly to the limitation of nitrogen and higher cell densities. Nitrogen limitation noticeably has been a key factor triggering sexual reproduction in almost all cells and subsequent resting spore formation. Other environmental factors include temperatures of 10 °C and 15 °C

but not at 20 °C and in cells with a valve diameter not more than 8 µm. Spores were observed to withstand anoxic conditions and low temperatures down to 2 °C. Vegetative cells arising from resting spores had an upper diameter of 14 µm.

Sexual reproduction starts initially with unequal division of cells within a vegetative chain of cells with a valve diameter of less than 8 µm (Fig. 1.6.1, a). The female gametangium or egg-producing cell is usually long and highly pigmented (Fig. 1.6.1, c) whereas the short male gametangium or spermatogonangium is only weakly pigmented (Fig. 1.6.1, c). Spermatogenesis follows two paths, in wider spermatogonangia (6–8 µm). The spermatogonangium then breaks open to release the two spermatogonangia (Fig. 1.6.1, e). Each spermatogonangia produces a quadriflagellate spermatogonangium, which in its turn divides into four uniflagellate spermatogonangia. If the original cells are narrower than 6 µm in valve diameter, then they produce either four or eight uniflagellate sperm cells. Conjugation of a sperm cell with an egg cell leads to the formation of a zygote which swells up to form an auxospore (Fig. 1.6.1, k). Fertilization involves bending of the egg cell to 45° angle (Fig. 1.6.1, d and f) where the sperm comes in contact with it. Later the sperm cell is drawn abruptly into the cell. Then the

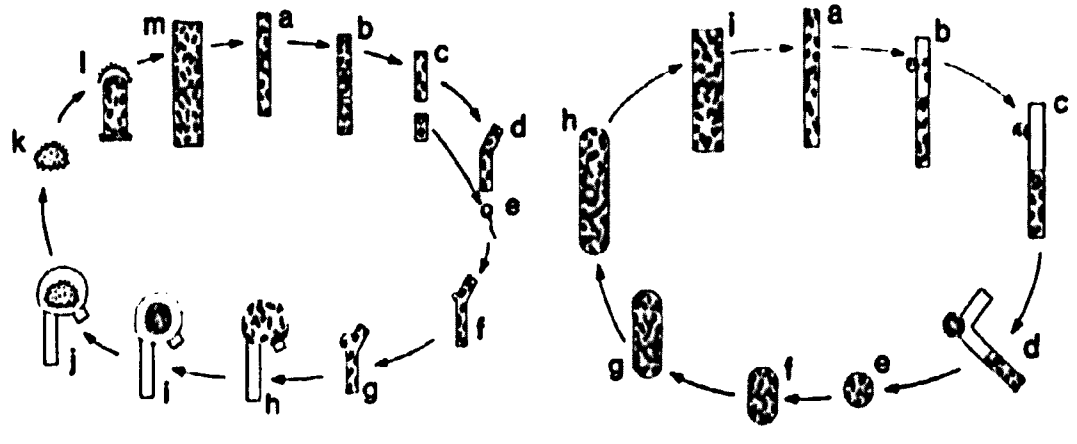


Fig. 1.6. Comparisons of the life cycles in *L. danicus* with or without resting-spore formation (1) cycle detailed in French and Hargraves (1985); sexuality (b-f) produces auxospore (h) within which the resting spore forms (j-k), later germinating (l-m) to produce cells of maximum diameter. Right cycle ("alternate ") exhibited by *L. danicus* var. *apora*; (a) Vegetative cells (narrow); (b) extrusion of cell contents, usually occurring midway along pervalvar axis (no sexuality); (c-d) extrusion completed, forming an auxospore-like structure; (e)auxospore like structure detached from parent cell; (f-g) elongation and germination; (h)

first cell division (presumed mitotic); (i) establishment of vegetative chain (diameter varies considerably) (From French and Hargraves 1986).

spines are formed (Fig. 1.6.1, i–k). This results in the formation of a heavily silicified and thick-walled auxospore from which a resting spore develops. Resting spore formation and its structure is another criteria used to discriminate *L. danicus* from *L. minimus*. Resting spores of *L. danicus* (Hargraves 1976) were described as, sub-spherical, constructed of two unequal valves, both bluntly and spiny. Whereas, the

1.6.1, i). Further to cellular contraction two silica walls are deposited and eventually resting spores of *L. minimus* (Hargraves 1990) vary in size from 5–6 µm in width and 7–10 µm in length, and with a shape of a spherical head and a cylindrical neck. Spores are usually highly ornamented with spines of varied size and shape. Its shape varies from hollow with depressed appearance to shorter and closed acute tips. At times, the spores may lack spines.

The alternate life cycle differs from that of the sexual life cycle in that resting spore are not formed. Instead, an auxospore-like structure of varying size is formed, into which the old cell content are extruded (Fig. 1.6.2, a–i). Later these structures undergo expansion and then normal division occurs.

1.6.2 ***Population dynamics.*** *Leptocylindrus danicus* is believed to be a widely distributed diatom ranging from polar to temperate and equatorial waters. The species is most commonly recorded in coastal water, but has been reported also in samples from offshore. In temperate regions, the species forms blooms during the warm periods of the year, but populations fluctuate in size, often forming blooms of lesser abundance. This means the species shows a year round distribution with different peaks of abundance. The understanding of the population distribution across a year comes mainly from the long-term stations at the Gulf of Naples (GoN) (Ribera d'Alcalà *et al.* 2004) and Narragansett Bay ((Ribera d'Alcalà *et al.* 2004, Karentz and Smayda 1998, Karentz and

Smayda 1984). Both sites are monitored weekly for phytoplankton abundance and environmental contextual data. The species shows the highest peaks of abundance during the spring and early summer (July–August) in the GoN when mean water temperature is around 22 °C whereas in Narragansett Bay, the peak occurs in May–June when mean water temperature is around 12 °C. This is huge variation within a species records. Apparently, other factors than temperature seem to govern the blooming of this species.

L. minimus shows a distribution range comparable to that of *L. danicus*, confined mostly to coastal ecosystem. However, *L. danicus* has a much larger distribution perhaps because of its warm water requirements. *L. danicus* and *L. minimus* recurrently co-occur, but the relative abundance varies greatly and their temporal isolation is also being recorded.

1.6.3 *Physiology.* *Leptocylindrus danicus* is found both in temperate and tropical waters. The species forms a major component of coastal blooms during the spring and summer. Photosynthesis and respiration are strongly influenced on temperature (Verity 1981a). Photosynthesis and cellular composition, growth and excretion of dissolved organic carbon (DOC) in *L. danicus* were found to have a temperature dependent day length effect with a Q_{10} value for photosynthesis of 7.0 (Verity 1981a). This means that at optimum temperature the species is very efficient in utilising low irradiances for photosynthesis. Optimum temperature of growth was reported between 15 °C to 20 °C, with maximum growth at 20 °C (Verity 1981a). Further increases in temperature lead to a decrease in growth. A higher irradiance period was found to promote growth with at least 12:12 L:D (Verity 1982a). Excretion rates were positively correlated with the photosynthetic rate and negatively correlated with chlorophyll *a* content (Verity 1981b). Dark (mitochondrial) respiration was reported to be dependent on previous light- and temperature history and this was more evident at higher temperatures (Verity 1982b).

Higher temperature increased respiration and day length affected oxygen consumption, but there seemed to be no clear effect on dark respiration (Verity 1982b).

1.7 The Gulf of Naples

The study site in the GoN, located on the western side of the Italian peninsula, opens to the Tyrrhenian Sea on the South-eastern side. The GoN can be divided into an oligotrophic offshore zone and a mesotrophic near-shore zone affected by ruff-off from the land (Pugnetti *et al.* 2006). The former is characterised by low chlorophyll concentrations in the surface water in the summer whereas the latter shows phytoplankton growth year-round. The institute, Stazione Zoologica Anton Dohrn Naples (SZN), since 1984 maintains a long-term ecological research station called MareChiara (MC) situated two nautical miles offshore in the transition zone between the coastal and open Tyrrhenian water. At the site, a bimodal distribution of phytoplankton species was observed taking into account the five most abundant species. Long term trends were also recorded for example there was decrease in cell size and an increase in abundances of diatoms and phytoflagellates (Ribera d'Alcalà *et al.* 2004).

1.8 Motivation for the study

Diatoms are major components of phytoplankton community in the coastal zones of the modern ocean. In coastal areas, planktonic diatoms are subject to fluctuations in their abundance due to the remarkable unevenness of their ecosystem at both spatial and temporal scales. Over the year, coastal phytoplankton species show peaks and sudden drops in abundances. Blooms in coastal areas are not single discrete events, but rather a series of fluctuations in which the biomass and the species composition change rapidly (i.e., species can be present in high numbers for a short period and later disappear completely from the water column; (Boero 1996, Cloern 1996). For example, in the GoN almost monospecific diatom blooms and rapid alternations of blooming species are

frequently observable features in seasons of high stratification (Zingone *et al.* 1990), (Ribera d'Alcalà *et al.* 2004). Subsequently, life cycle phase transitions can have a major effect on succession patterns (Hansson 1995, McQuoid and Hobson 1995, Rengefors and Anderson 1998). Hence, the factors regulating the seasonal succession and bloom events of marine phytoplanktons, especially harmful algal bloom species, and species that are major contributors of algal blooms, are of great interest both from an ecological perspective and for coastal management (Anderson and Rengefors 2006). Despite years of efforts in monitoring and experimental studies, patterns of succession and bloom events remain largely unpredictable (Anderson and Rengefors 2006). To date, efforts have been concentrated on the study of the mechanisms underlying the spring bloom and on the possible prediction of its timing, whereas there is a poor understanding of the factors that drive individual species seasonal fluctuation (Zingone and Wyatt 2005) and shape their range. The most accepted paradigm for phytoplankton variability at a seasonal scale is that individual species have traits that allow them to thrive under particular environmental conditions (Margalef 1978, Reynolds 1998, Kneitel and Chase 2004). In this view, the so-called habitat template hypothesis (Reynolds 2001, Smayda 2000), the growth of species sharing similar ecophysiological traits is stimulated by the regular occurrence of a set of environmental conditions that recur yearly over the seasonal time (Kneitel and Chase 2004). On the other hand there are multiple evidences of species reoccurring in certain periods of the year despite marked interannual variations in environmental conditions (Ribera d'Alcalà *et al.* 2004). Eilertsen *et al.* (1995) provided some evidence for photoperiodic control of blooms in diatoms. While some species increase their abundance in restricted periods of the year, other species show multiple peaks over the year, in seasons which markedly differ in terms of physical and chemical parameters (McDonald *et al.* 2007); still others show relatively long bloom times (months), encompassing a wide range of environmental conditions (Ribera d'Alcalà *et al.* 2004).

1.9 Goals and thesis framework

During the research the major question being addressed is: what are the characteristics of the species that favour it to grow under a wide range of environmental conditions? Current views on phytoplankton species are that coastal species, and especially those that are widespread and abundant, are very plastic from the physiological point of view and generally have a wide tolerance for environmental parameter variability. An alternative hypothesis to explain the wide temporal window of occurrence of *L. danicus* in the GoN is that there are different species or genetically distinct populations under this name that alternate over the year, producing blooms under different environmental conditions. The two hypotheses however are not mutually exclusive, as there could be different genotypes over the year each with high physiological plasticity. Together or alternatively, these **two aspects of species biology determine its ability to bloom** under varied niche conditions.

The main objective of this thesis was to obtain a better understanding of the species diversity in the genus and biogeography of the marine diatom genus *Leptocylindrus*. Specifically investigated are;

1) Whether *L. danicus* is a single species or comprises multiple (pseudo)-cryptic species, by analysing variation patterns in selected phylogenetic markers, life cycle patterns and morphological characteristics,

2) Whether the distribution of *L. danicus* is truly cosmopolitan or whether potential (pseudo)cryptic entities show different, possible restricted, geographic distributions; and

3) Whether the different biological entities have a different physiological adaptability alternating across different seasons.

In the first part of this thesis, the diversity of the species *L. danicus,* life cycle patterns and distribution of the different entities previously recognised within the species is elucidated.

In **Chapter II** the morphology and phylogenetic relationships of the five species discovered within the genus *Leptocylindrus* has been described. Sexual reproduction was successfully initiated experimentally and its characteristics, together with structure of resting spores and alternate life cycle patterns are also described.

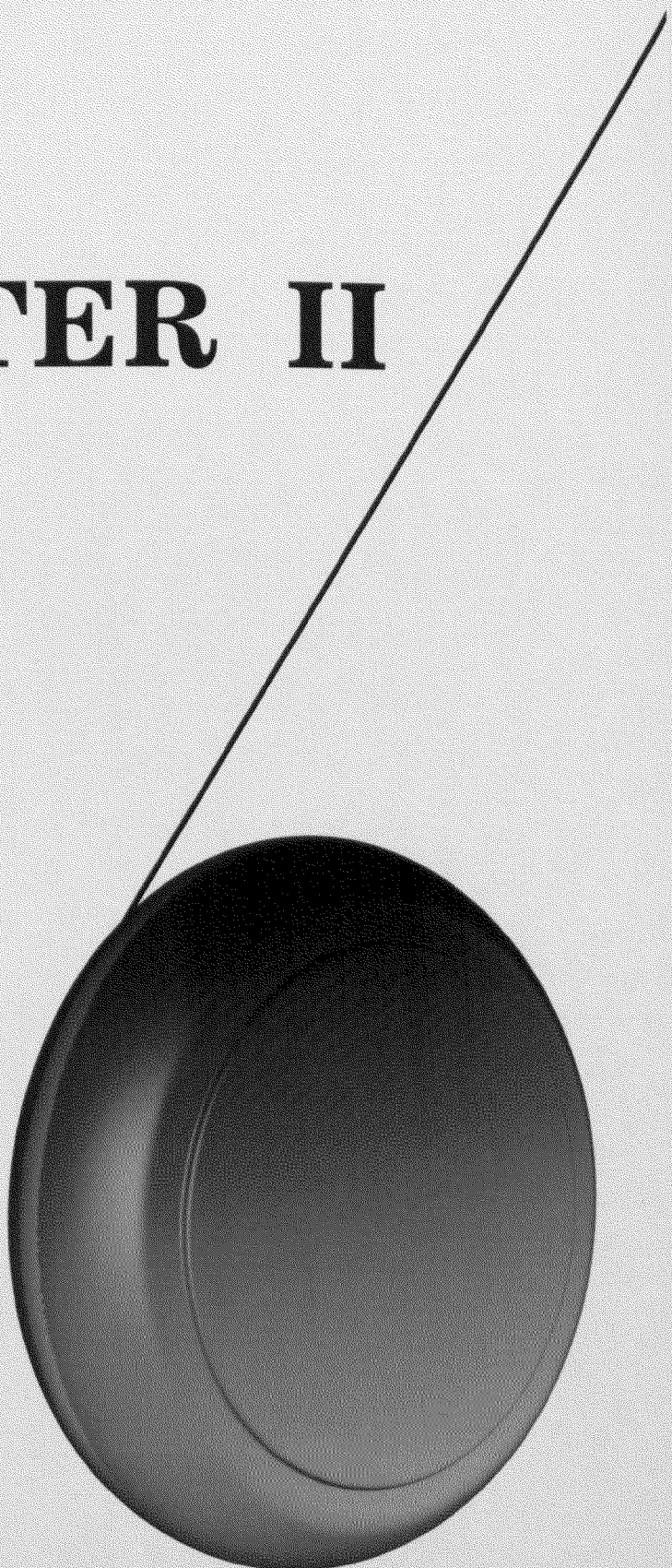
In **Chapter III** the distribution of *Leptocylindrus* species based on the databases of environmental sequences including BioMarKs (European database for metagenomics), *Tara* Oceans database (the global database) and GenBank (NCBI) has been described. These databases were explored for the six species described in Chapter II together with any other sequences grouping with them to understand the biogeography and seasonality of the species.

In the second part of this thesis the biochemistry of the six species was studied. Further, the physiology of *L. danicus* and *L. aporus* were studied, which were the two most abundant species in the GoN with opposite seasonality.

In **Chapter IV** the importance and ecological relevance of oxylipins and the different profiles of the five discovered species has been described. The diversity in the metabolites production helps structure the ocean ecosystem.

In **Chapter V** the physiological differences of the species *L. danicus* and *L. aporus* in response to variation in temperature has been described. Even closely related species can differ in functional properties such as production of secondary metabolites and this has implications in the functioning of the ecosystem.

CHAPTER II



Morphological and molecular diversity of *Leptocylindraceae* species from the Gulf of Naples

2.1 Introduction

Since the description of the first diatom, *Bacillaria paradoxa* Gmelin L.M. (1791), the number of know species in this group of marine algae has incessantly increased, with brisk upsurges following the introduction of new methodologies. Until the 1960's, diatomists observed diatoms almost exclusively using light microscopy (LM). This tool allowed studying frustule shape as well as features of the living cell, thus leading over time to the description of numerous species. With the onset of transmission and scanning electron microscopy (TEM, SEM) frustule ultrastructure was revealed in all its fine details, and as a consequence a wealth of additional, morphological diverse species was discovered. In the 1980s, PCR amplification of marker DNA sequences demonstrated yet another level of diversity: cryptic and semi-cryptic diversity, i.e., morphological taxa consisting of genetically distinct but morphologically indistinguishable or barely distinguishable species. In cases where this was tested, the cryptic species were usually also reproductively isolated (e.g. Amato et al. 2007, Lundholm et al. 2012). At present, ca. 200,000 diatom species are estimated to exist (Mann and Droop 1996) of which about 10,000 have been described (Mann *et al.* 1999). The wealth of DNA sequences generated from specimens from all over the diatom diversity has also permitted reconstruction of detailed phylogenies. Results show that the pennate diatoms form a clade inside a grade of centric diatoms. Pennates have usually elongated valves, with a distinctive keel, called a midrib, with rows of pores perpendicular upon it; centric diatoms generally have valves that are circular or have a multi-polar outline, with a central ring or area from which rows of pores radiate (Kooistra *et al.* 2007).

Results of recent multidisciplinary taxonomic studies within common diatom genera uncovered considerable diversity in the pennate genera *Pseudo-nitzschia* (Lundholm et al. 2012 and literature therein) and *Sellaphora* (Evans *et al.* 2008) and in the centric species *Cyclotella meneghiniana* (Beszteri *et al.* 2007) and *Skeletonema* (Sarno *et al.* 2005, Sarno *et al.* 2007, Zingone *et al.* 2005, Kooistra *et al.* 2008). The highly diverse centric genus *Chaetoceros* (Rines and Hargraves 1990) also shows marked diversity within morphologically perceived species (Kooistra *et al.* 2010). Cryptic and semi-cryptic species in all these genera are genetically closely related, showing only a marked differentiation in rapidly evolving genetic markers such as the nuclear large subunit or the internal transcribed spacers of the nuclear ribosomal RNA cistron. In contrast, other diatom genera may be species-poor. Phylogenies reveal an apparent paucity of species diversity within radial centric genera, which are the oldest and most basal of diatoms in the phylogeny and also the earliest ones appearing in the fossil record (Kooistra *et al.* 2007). This paucity could result from limited taxonomic work focusing on this group, despite the important role of many radial centric genera as primary producers in the oceanic or coastal plankton. Alternatively, the apparent paucity of species in these genera could be real, with each of these genera consisting of just a few genetically distinct survivor species of a once far more extensive diversity (Kooistra *et al.* 2007).

Among apparently ancient diatom lineages, the genus *Leptocylindrus* merits detailed exploration of its ultrastructural and genetic diversity. Its species consist of cylindrical cells with two or more plastids and valves with a simple morphology. Molecular phylogenies usually resolve the genus as sister lineage to a clade containing all other diatoms. Currently, two main species are widely recognized; *L. danicus* (Cleve 1889), within which the variety *L. danicus* var. *apora* Hargraves & French (1986) was described, and *L. minimus* Gran (1915; see also (Hargraves 1990, Rivera *et al.* 2002). The taxonomic affiliation of the species *L. mediterraneus* (H. Peragallo) Hasle has been long questioned (Hasle and Syvertsen 1997, Gomez 2007). In fact, it is even doubted

that this species is a diatom, as only empty frustules with a peculiar ultrastructure have been observed in nature, almost exclusively observed as colonized by the protozoan *Solenicola setigera* Pavillard (Gomez 2007). Two other species described in the past, namely *L. belgicus* Meunier (1915) and *L. adriaticus* Schröder (1908), have subsequently been considered respectively as a synonym of *L. minimus* (Hustedt 1962, Hendey 1964), and a variety of *L. danicus* (Schiller 1929). Finally, the description of *L. curvatus* Skvortzov (1931), which is based on a drawing of undulating chains of cells with small plastids, has hardly been followed by any report of the species, whose identity remains uncertain. So far, only rarely have *L. danicus* and *L. minimus* been investigated from the molecular point of view, and one SSU DNA sequence for each of these species has been deposited in GenBank (AJ535175, AJ535176).

The life cycle of *Leptocylindrus* species shows a wide diversity. Formation of highly silicified resting spores in *L. danicus* was first described by Murray and Hjort (1912), whereas Gran (1915) first reported that these resting spores develop directly from auxospores. Subsequently, Davis et al. (1980) and French and Hargraves (1985, 1986) demonstrated the formation of spores in *in vitro* cultures, confirming that *L. danicus* is one of the few diatom species forming these stages following sexual reproduction. In contrast, the variety *L. danicus* var. *aporus* restores its cell size by asexual autoenlargement (Hargraves & French 1986). Instead *L. minimus* forms resting stages during the vegetative growth, which is similar to other diatoms (Hargraves, 1990).

Leptocylindrus danicus has been reported throughout the world's oceans, except in Arctic waters, and forms major blooms in coastal waters. *Leptocylindrus minimus* is widespread in the Mediterranean Sea, along the North Atlantic coasts of Europe and America as well as in the eastern and western South Pacific. In the latter areas, the species is enumerated in the list of harmful species, since its blooms may cause damages to fish (Clement and Lembeye 1993, Hallegraeff *et al.* 2003). In the Gulf of Naples (GoN), specimens assignable to *L. danicus* are observed throughout the year, but

show two distinct bloom phases, a conspicuous one in summer and a more modest one in the fall. *Leptocylindrus minimus* is recurrently found in autumn.

The aim of this study is to investigate the diversity of the genus *Leptocylindrus* in the GoN by combining morphological and molecular data with information on the life cycle. To this end, a series of strains were isolated from the GoN over an entire seasonal cycle. Strains were examined using six molecular markers as well as light- and electron-microscopic observations. In addition, to elucidate the differences in the life cycle patterns among the taxa investigated, spore induction studies were conducted. Attempts were made to link morphological characteristics of individual taxa to the descriptions of already known species. First, sequences of each marker were sorted into (near) identical groups. Within each of these groups three randomly picked strains were checked if they shared highly similar ultrastructural details of the cell wall and similar plastid shapes. Subsequently the groups of specimens were verified if they could be assigned to known species descriptions or if they belonged to species new to science. Second, phylogenies were reconstructed to assess if the genus *Leptocylindrus* is monophyletic and if the taxonomic units are phylogenetically closely related or not. Based on molecular and morphological results, it is proposed to raise the taxon *L. danicus* var. *apora* to the rank of species and two new species, *L. hargravesii* and *L. convexus* are described. In addition, based on morphological and phylogenetic criteria, it is suggested to establish the new genus *Tenuicylindrus* including a single species *T. belgicus* based on the description of *L. belgicus*, so far considered as a heterotypic synonym of *Leptocylindrus minimus*.

2.2 Methods

2.2.1 Strain isolation. 86 strains were obtained to assess the morphological and genetic diversity of *Leptocylindrus* (Table 2.1). Single cells or chains were gathered from net-samples collected at the LTER MareChiara (LTER-MC) in the GoN (40.80° N, 14.25°

Table 2.1. Strains used in the analysis. All strains are from the Gulf of Naples, except CCMP 1856 which was isolated by Hargraves, P., from the Gulf of Mexico.

Strain Information	Collection Date	SSU rDNA	LSU rDNA	ITS	Chl. SSU rDNA	psbC	rbcL
Leptocylindrus aporus							
SZN-B702	-/10/2009	+	+	+	+	+	+
SZN-B703	-/10/2009	+	+	+	+	+	+
SZN-B704♦	-/10/2009	+	+	+	+	+	+
SZN-B723	03/08/2010			+			
SZN-B724	03/08/2010			+			
SZN-B725	03/08/2010			+			
SZN-B726	03/08/2010			+			
SZN-B727	03/08/2010		+	+			
SZN-B728	03/08/2010			+			
SZN-B729	10/08/2010			+			
SZN-B730	10/08/2010			+			
SZN-B731	10/08/2010			+			
SZN-B732	10/08/2010		+	+			
SZN-B733	10/08/2010			+			
SZN-B734	10/08/2010			+			
SZN-B735	14/08/2010			+			
SZN-B736	14/08/2010			+			
SZN-B737	14/08/2010			+			
SZN-B738	14/08/2010			+			
SZN-B651♠♦	21/08/2010	+	+	+	+	+	+
SZN-B743♦	12/10/2010	+	+	+	+	+	+
SZN-B744	19/10/2010			+			+
SZN-B745	19/10/2010			+			
SZN-B746	19/10/2010			+			
SZN-B747	19/10/2010			+			
SZN-B748	19/10/2010			+			
SZN-B749	19/10/2010			+			
SZN-B750	19/10/2010			+			
SZN-B751	19/10/2010			+			
SZN-B752	19/10/2010			+			+
SZN-B753♠	19/10/2010			+			
SZN-B754	19/10/2010			+			
SZN-B758	27/10/2010			+			
SZN-B759	27/10/2010			+			
SZN-B760	02/11/2010			+			
SZN-B761	02/11/2010			+			
SZN-B762	02/11/2010			+			
SZN-B763	18/11/2010			+			
SZN-B764	18/11/2010			+			
SZN-B782	25/01/2011			+			
SZN-B784	25/01/2011			+			
Leptocylindrus convexus							
SZN-B709	09/03/2010	+	+	+	+	+	+
SZN-B768♠♦	21/12/2010	+	+	+	+	+	+
SZN-B769	21/12/2010			+			
SZN-B770♦	21/12/2010			+			
SZN-B771♦	21/12/2010			+			
SZN-B774	25/01/2011			+			
SZN-B775	25/01/2011	+	+	+	+	+	Contd...

SZN-B776	25/01/2011	+	+	+	+	+	+
SZN-B777	25/01/2011			+			
SZN-B778♠	25/01/2011			+			
SZN-B779	25/01/2011			+			
SZN-B780	25/01/2011			+			
SZN-B783	25/01/2011			+			
SZN-B788	08/02/2011			+			
Leptocylindrus danicus							
SZN-B705	15/02/2010	+	+	+	+	+	+
SZN-B706	15/02/2010	+	+	+	+	+	+
SZN-B707♠♦	15/02/2010	+	+	+	+	+	+
SZN-B708	15/02/2010	+	+	+	+	+	+
SZN-B710	30/03/2010		+	+			
SZN-B711♦	21/04/2010		+	+		+	
SZN-B712	21/04/2010		+	+			
SZN-B713	21/04/2010		+	+			
SZN-B714	21/04/2010		+	+			
SZN-B715	15/06/2010		+	+			
SZN-B716	15/06/2010		+	+			
SZN-B717	15/06/2010		+	+			
SZN-B650♠♦	15/06/2010	+	+	+	+	+	+
SZN-B718	13/07/2010			+			
SZN-B765	18/11/2010			+			
SZN-B786	25/01/2011			+			
SZN-B789	15/02/2010			+			
SZN-B790	15/02/2010			+			
SZN-B791	15/02/2010			+			
Leptocylindrus hargravesii							
SZN-B772♠♦	21/12/2010	+	+	+	+	+	+
SZN-B773	21/12/2010	+	+	+	+	+	+
SZN-B781♠♦	25/01/2011	+	+	+	+	+	+
CCMP 1856		+	+	+	+	+	+
Tenuicylindrus belgicus							
SZN-B739♠♦	02/10/2010	+	+	+	+	+	+
SZN-B740	02/10/2010		+	+			
SZN-B741	02/10/2010		+	+			
SZN-B742	02/10/2010		+	+			
SZN-B755♦	19/10/2010	+	+	+	+	+	+
SZN-B756♠♦	19/10/2010	+	+	+	+	+	+
SZN-B757	19/10/2010		+	+			

♠ Strains used for life cycle experiments

♦ Strains used for SEM and TEM observations

E) from February 2010 to February 2011. Strain CCMP 1856 was obtained from the National Centre for Micro Algae and Microbiota (NCMA; formerly CCMP).

The procedure for cell isolation was as follows: the net was towed gently through the surface layer of the water column (about 1 to 2 m), and the obtained plankton sample

was stored at about 15 °C in the dark until return to the laboratory. Small amounts (ca. 0.5 to 1 ml) of the sample were placed in single wells of a tissue culture plate and diluted with filtered seawater. Under the LM, *L. danicus* specimens were recognized based on cell morphology. Single cells or short chains of *L. danicus* were isolated with drawn-out Pasteur pipettes under an inverted light microscope, washed in sterile seawater and placed in wells of a tissue culture plate containing 2 ml K growth medium (Keller *et al.* 1987). The culture plates were incubated at 20 °C, with a photon fluency rate of 60 µmol photons $m^{-2}s^{-1}$ (provided by cool-white fluorescent tubes) in a 12:12 h light: dark (L:D) photo cycle for one week to allow the growth of isolated cells. Strains were screened for purity and, if unialgal, used as starter culture for further growth in 70ml polystyrene flasks containing 25 ml of sterile K medium. Fully grown cultures are further used for microscopic observations and molecular characterization.

2.2.2 Microscopic observations. Microscopic observations were made on selected strains including representatives of each genetically distinct group. Morphological features were observed in LM, TEM and SEM. All LM observations were carried out on exponentially growing cultures and natural samples using a Zeiss Axiophot microscope (Carl Zeiss, Oberkochen, Germany) equipped with Nomarski differential interference contrast (DIC), phase contrast, and bright-field optics. Light micrographs were taken using a Zeiss Axiocam digital camera (Carl Zeiss, Oberkochen, Germany).

Samples collected for electron microscopy were fixed immediately with 40% formalin at a concentration of 4% (V/V) and was mixed thoroughly with samples, this, in addition to preserve the cells, will help in dehydration of cells. For preparation of sample for TEM observation exponentially growing cultures were treated with 10% H_2SO_4 and 10% HNO_3, gently boiled and washed with distilled water for few times to remove the residual acid. Acid treatment helps in the removal of organic material and detaching the silica frustule, but leaves them intact. Acid cleaned material was mounted

on Formvar-coated grids and examined under a TEM LEO 912AB. For preparation of samples for SEM observation cultures were prefixed either with formalin or Osmium. A small aliquot of preserved culture was filtered through nylon membrane filter, taking care to have well distributed cells on the filter. An increasing percentage of ethanol (25%, 50%, 60%, 75%, 90% and 100%) was passed through the filter to completely dehydrate the cells. The dehydrated filter with cells was passed through critical point drying, splutter coated with gold-palladium and mounted on a stab. Cells were examined using a JEOL JSM-6500F SEM (JEOL-USA Inc., Peabody, MA, USA).

2.2.3 DNA extraction, PCR amplification, sequencing, and phylogenetic analyses. Extraction of total DNA from individual strains was performed using a rapid, simple and efficient protocol. Exponentially growing cultures (density of 50,000 cell ml^{-1}) were used for DNA extraction. Cells from 2 ml culture were harvested by centrifugation at 10,000 *g* for 10 min at 15 °C. The cell pellet was then suspended in 400 µl extraction buffer (100 mM Tris HCL pH 8.0, 50 mM EDTA, 500 mM NaCl, 1.5% SDS, 0.2% β-mercaptoethanol) containing 2 µl of RNase (10 mg µl-1, Roche S.p.A., Milan, Italy). The mixture was incubated at 65 °C in a water bath for 20 min, vortexing every 5 min to allow efficient disruption of cells and release of cell contents. After incubation, 100 µl of 5 M potassium acetate was added and mixed well by inverting the tubes. This precipitates the cell debris leaving DNA and other soluble molecules in the supernatant. The supernatant (500 µl) was collected by centrifuging at 10,000 *g* for 10 min at room temperature. To the collected supernatant 500 µl ice cold isopropanol and 100 µl of 3 M sodium acetate was added to precipitate DNA. The precipitated DNA was pelleted by centrifugation at 10,000 *g* for 10 min at room temperature. Then the obtained DNA pellet was washed with 250 µl of 75% ethanol to remove salts. Residual ethanol was removed by drying the DNA pellet in a vacuum drier for 8 min at 60 °C. The clean dry

DNA pellet was resuspended in 0.1x TE buffer (1x TE buffer: 10 mM Tris-HCl, 1 mM EDTA) and stored at 4 °C for short term use or at -20 °C for long term use.

The following markers were amplified. Of the nuclear ribosomal RNA-coding cistron were amplified the SSU (18S) rDNA, the Internal Transcribed Spacer region (ITS1, 5.8S rDNA and ITS2) and ca. 700 bp of the 5′-end of the LSU rDNA; of the plastid DNA were amplified the SSU (16S) rDNA, the Rubisco-large subunit (*rbc*L) and the photo-system binding complex protein coding gene (*psb*C) following amplification using primers listed in Table 2.2 and Polymerase Chain Reaction (PCR) protocols present in Table 2.3. The PCR products were purified using a QIAquick gel extraction kit (Qiagen, Milan, Italy) according to the manufacturer's instructions. Purified products were analysed on an automated Capillary Electrophoresis Sequencer "3730 DNA Analyzer" (Applied Biosystems, CA, U.S.A.) in collaboration with the Sevizio Biologia Moleculare at the institute. Sequencing primers were the same as those utilized for amplification. In addition, internal sequencing primers were deployed for the products of the nuclear and plastid SSU rDNA, the *rbc*L and the *psb*C, as these products were too long to be read with amplification primers alone. Internal sequencing primers are also listed in Table 2.2.

The obtained sequences were aligned (along with sequences from *Bolidomonas* spp. and other outgroup diatom sequences) using Bioedit v. 7.1.3 (http://www.mbio.ncsu.edu-/bioedit/page2.html) and then adjusted by eyeball in the sequence alignment editor Se-Al version 2.0a11 (Rambaut 1996-2002). In addition, the DNA sequences of different strains were aligned by Clustal-W alignment. The secondary structure of ITS regions was predicted using the *RNAfold* webserver (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) and used for calculating the number of compensatory base changes (CBC). Helices were compared with each other to identify possible homologous regions to assist alignment. Highly similar sequences were obtained from GenBank using BLAST searches. Prior to phylogenetic analysis, identical sequences were removed

Table 2.2. List of primers used for the amplification of different regions with the specific annealing and extension temperature with time of hold.

Gene Region	Primer Name	Sequence	Purpose	Product (bp)	Annealing Temp (°C)	Annealing Time (min)	Extension Temp (°C)	Extension Time (min)
SSU rDNA	SSUF	AACCTGGTTGATCCTGCCAGT	PCR/Seq F	~1800	62	0.5	72	**2**
	SSUR	TGATCCTTCTGCAGGTTCACCTAC	PCR/Seq R					
	528F	GCGGTAATTCCAGCTCCAA	Seq F	~600				
	1055R	ACGGCCATGCACCACCACCCAT	Seq R					
LSU rDNA	D1R	ACCCGCTGAATTTAAGCATA	PCR/Seq F	~750	55	1	72	**1**
	D2R	TGAAAAGGACTTTGAAAAGA	PCR/Seq R					
ITS rDNA	ITS1	TCCGTAGGTGAACCTGCGG	PCR/Seq F		46	1	72	**2**
	ITS4	TCCTCCGCTTATTGATATGC	PCR/Seq R					
	LDF 1*	TACGTCCCTGCCCTTTGTAC	PCR/Seq F	~750	60	1	72	**1**
	LDR 4*	GGGGGTGTCACCCTCTATG	PCR/Seq R					
rbcL	DPrbcL1	AAGGAGAAATHAATGTCT	PCR/Seq F	~1470	55	1	72	**2**
	DPrbcL7	AARCAACCTTGTGTAAGTCTC	PCR/Seq R					
	rbcLI1F	TTAGAAGACATGCGTATT	Seq F					
	rbcLI1R	CAGTGTAACCCATAAC	Seq R					
psbC	Fo = psbC+	CACGACCWGAATGCCACCAAT	PCR/Seq F	~1100	55	1	72	**2**
	Ro = psbC-	ACAGGMTTYGCTTGGTGGAGTGG	PCR/Seq R					
	F1 = psbC221+	ACGCATTGTTTCACCACC	Seq F					
	F2 = psbC499+	ACGTGCCCAAGAGAATGGTTTTG	Seq F					
	R1 = psbC857-	CTTTGGTTATGACTGGCGTG	Seq R					
	R2 = psbC587-	ATCTTGTTGGTGGTCATATTTGG	Seq R					
Plastid SSU rDNA	PLA491F	GAGGAATAAGCATCGGCTAA	PCR/Seq F	~800	62	0.5	72	**1**
	OXY1313R	**CTTCAYGYAGGCGAGTTGCAGC**	**PCR/Seq R**					

* *Leptocylindrus aporus* specific primers were designed and used for the amplification of ITS region of rDNA

Table 2.3. General PCR protocol used for the amplification of different nuclear and chloroplast genes for the phylogenetic analysis.

General PCR Protocol			
Step	Temp (°C)	Duration (min)	No. of cycles
Initial denaturation	94	4 min	1
Denaturation	94	30 sec	35
Annealing	S*	S*	
Extension	72	S*	
Final extension	72	10 min	1
Final Hold	**10**	∞	**1**

S* specific temperature or time of hold has been mentioned in Table 2.2.

from the alignments until only pairs of identical sequences remained. Maximum likelihood and maximum parsimony trees were constructed utilizing (Phylogenetic Analyses Using Parsimony; PAUP* version 4.0 and other methods) (Swofford 1998). Most parsimonious trees were obtained using heuristic searches. Branches were collapsed if their minimum length was zero; character state optimalisation was done using accelerated transformation, multistate taxa were treated as polymorphisms and gaps, if occurring were treated as missing data. Heuristic searches were carried out by keeping best trees only; trees were started by random step-wise addition and tree-bisection-reconnection branch-swapping, performing ten replicate runs. Bootstrap values associated to internodes were based on 1000 bootstrap replicates; each replicate was carried out as described for heuristic searchers with a single run per bootstrap replicate.

Maximum likelihood trees were obtained as follows. Modeltest version 3.06 (Posada and Crandall 2001) was used to select optimal base substitution models and values for base composition, % invariable sites, and gamma shape parameter according to the Akaike information criterion (AIC). Maximum likelihood (ML) analyses of the alignments were performed under the full heuristic search option in PAUP* version 4.0b10 and were constrained with values obtained by Modeltest. Only best trees were

kept, trees were started by random step-wise addition and tree-bisection-reconnection branch-swapping, performing ten replicate runs. Bootstrap values associated to internodes were based on 1000 bootstrap replicates; each replicate was carried out as described for heuristic searchers with a single run per bootstrap replicate.

2.2.4 Spore induction. For an explicit description of the identity of different species it is important to elucidate the differences in the life cycle patterns, in particular the sexual reproductive stage. Spore induction studies were done with two strains representative of each genetically distinct clade. Sexual reproduction experiments were carried out as described in French and Hargraves (French and Hargraves 1980) by simple reduction of the nutrient levels in the medium, designated as "T" medium (containing ammonia at 15 µM; phosphate at 7 µM; silicate at 50 µM; trace metals, vitamins, and Trizma as in f/2) (Guillard and Ryther 1962), with at a slightly lower temperature of 16 °C using cool fluorescent light of 80 µmol photons $m^{-2}s^{-1}$ intensity and a 12:12 L:D cycle.

The strains submitted to these reduced nutrient conditions (Table 2.1) were observed regularly in the LM for 20 days. On cultures where spore formation was observed, pictures were taken in the light microscope and samples were prepared for SEM observations as described above.

2.3 Results

Five distinct species were recognized in the material mainly from the GoN and also including the only strain available in a culture collection. In the following a morphological description and a taxonomic assignment is presented for each of the 5 species. The diagnoses of already known species were emended to include new observations and make them comparable among the species.

2.3.1. ***Leptocylindrus aporus*** (French and Hargraves 1986) Nanjappa et Zingone, *comb. nov.* (Fig. 2.1, A–H)

Syn: ***Leptocylindrus danicus*** **var.** ***apora*** **Hargraves**

Emended diagnosis: Cells cylindrical, 3.5–10.6 μm in diameter, 12.5–33 μm in pervalvar length, solitary or forming short filamentous chains. Plastids few to many, generally ovoidal. Valves slightly convex or concave, with a central annulus delimiting a group of poroids. No distinct sub-central pore. Poroid areolae (10–14 areolae in 1 μm) in radiating striae (10–23 striae in 1 μm) Mantle curved, with poroid areolae in parallel striae (8–10 striae in 1 μm). Short triangular, often blunt spines along a ring at the margin between valve face and mantle. Girdle consisting of elongated and nearly trapezoidal half bands, with irregular rows of areolae along their short axis. Resting spores not observed.

Holotype: Figure 6 J in Hargraves and French 1986

Isotype: A permanent slide of strain SZN-B650 deposited in the SZN Museum as no. SZN-B650-01.

Iconotype: Figure 2.1, A–H

Materials examined: Strains SZN-B650, SZN-B707 and SZN-B714.

Type locality: Narragansett Bay, Rhode island, USA (41°30.3'N, 71°25.2'W)

Description. The cells are cylindrical, 4–7.5 μm in material from the GoN, and joined by their faces (Fig. 2.1, A–D, and Table 2.4) to form filamentous colonies of up to 25 cells. Solitary cells are also often found in both culture and nature. Each cell contains 3–13 plastids (even one or two in natural samples during summer), which are ovoid and rarely discoid or elongated and are distributed along the periphery of the cell (Fig. 2.1, A–C). The valve faces are slightly convex or concave, and include a slightly denser central area with a group of areolae delimited by a thickened hyaline annulus (Fig. 2.1, E and G), from which striae of areolae (10–14 in 1 μm, Table 2.5) radiate towards the

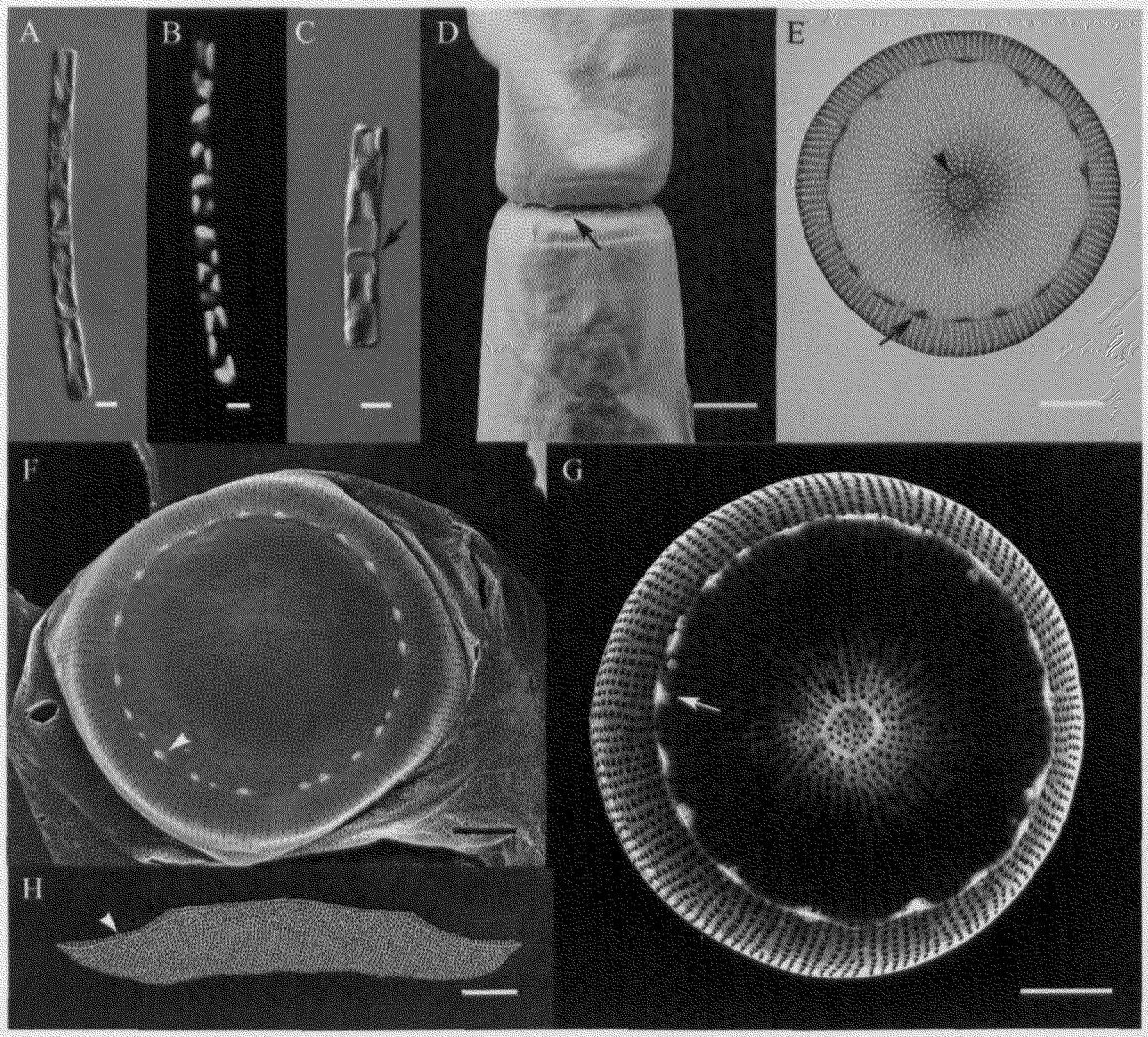


Fig. 2.1. *Leptocylindrus aporus*, strain SZN-B651 (A–D, F–H) and strain SZN-B743 (E). LM, DIC: A and C; epifluorescence: B. SEM: D and F. TEM: G, E and H. (A) Colony of four cells. Scale bar, 5 µm. (B) Same colony as in A, note the shape of the plastids. Scale bar, 5 µm. (C) Colony of two cells, note the shape of the joining (arrow). Scale bar, 5 µm. (D) Detail of two sibling cells in lateral view, joined by the valve face, with flaps alternating along the marginal ring (arrow). (E) Valve with central annulus (arrowhead) and blunt triangular marginal flaps (arrow); note the absence of the sub-central pore. Scale bar, 1 µm. (F) Cell in valve view, with flaps (arrowhead) at the boundary between valve face and mantle. Scale bar, 1 µm. (G) Valve with central annulus (arrowhead) and blunt triangular marginal flaps (arrow); note the absence of the sub-central pore. Scale bar, 1 µm. (H) Copula with a trapezoidal shape and an almost continuous hyaline line close to one border (arrowhead). Scale bar, 1 µm.

Table 2.4. Light microscope morphometric characters in *Leptocylindrus* and *Tenuicylindrus* species.

		L. aporus	*L. convexsus*	*L. danicus*	*L. hargravesii*	*T. belgicus*
Cell diameter (μm)	min–max	4–7.5	3–8	3–13	3–15	2–2
	mean ± SD	5.4 ± 1.38	6 ± 1.52	7.2 ± 2.53	9.3 ± 4.54	2 ± 0
	n	50	55	50	40	40
Perivalvar length (μm)	min–max	12.5–33	22–65	22–75	30–90	23.6–50
	mean ± SD	21.9 ± 5.37	40.1 ± 9.54	35 ± 10.11	54.8 ± 14.95	36.1 ± 7.16
	n	50	55	50	40	40
Chloroplast number	min–max	3–13	3–11	7–36	9–55	2–2
	mean ± SD	7 ± 2.6	7.9 ± 3.43	15.6 ± 6.34	25.6 ± 14.22	2 ± 0
	n	50	55	50	40	40
Chloroplast size	min–max	2–6	2 – 14	2–6	2–6	7–17.8
	mean ± SD	3.21 ± 1.0	4.1 ± 2.82	2.3 ± 0.61	3 ± 1.33	10.91 ± 2.39
	n	70	100	80	166	80
Cells per colony	min–max	2–24	2–68	2–165	2–162	2–14
	mean ± SD	7.52 ± 4.5	8.5 ± 21.25	39.22 ± 34.1	20.31 ± 29.84	4.1 ± 1.71
	n	100	100	100	100	153

Table 2.5. Electron microscope morphometric characters in *Leptocylindrus* and *Tenuicylindrus* species.

			L. aporus	*L. convexus*	*L. danicus*	*L. hargravesii*	*T. belgicus*
Cell	**Cell diameter**	min–max	5.00–7.44	4.54–7.70	4.53–7.40	4.30–15.42	1.30–2.10
		mean ± SD	5.95 ± 0.73	6.84 ± 0.85	5.82 ± 7.40	10.52 ± 4.43	1.71 ± 0.25
		N	14	16	9	16	16
Valve face	**Valve face diameter (µm)**	min–max	3.04–5.04	3.03–5.00	3.65–6.52	3.04–12.22	1.11–1.96
		mean ± SD	4.26 ± 0.52	4.51 ± 0.56	4.77 ± 0.83	8.77 ± 3.64	1.54 ± 0.22
		N	14	16	9	16	16
	Valve face to mantle ratio	min–max	2.90–8.40	2.40–4.04	5.60–11.45	5.30–14.30	
		mean ± SD	5.32 ± 1.52	3.20 ± 0.42	7.40 ±1.93	8.50 ± 2.51	
		N	13	14	7	16	
	Striae (1 µm, Centre)	min–max	10–23	10–14	11–16	6–11	13–18
		mean ± SD	15.36 ± 4.25	11.80 ± 1.15	12.60 ± 1.95	8.38 ± 1.26	15.77 ± 1.36
		N	14	16	5	16	13
	Striae (1 µm, End)	min–max	9–13	6–10	8–13	7–11	7–11
		mean ± SD	10.21 ± 1.31	8.19 ± 1.17	10.38 ± 1.51	8.40 ± 1.06	8.71 ± 1.14
		N	14	16	8	16	14
	Areolae no. (1 µm)	min–max	10–14	10–14	18–30	10–14	12–22
		mean ± SD	12.57 ± 1.22	11.91 ± 1.13	21.50 ± 4.00	12.31 ± 1.01	15.71 ± 3.29
		N	14	16	8	16	14
	Radial areolae diameter (µm)	min–max	0.02–0.035	0.02–0.04	0.014–0.03	0.02–0.055	0.02–0.03
		mean ± SD	0.03 ± 0.001	0.03 ± 0.005	0.025 ± 0.005	0.032 ± 0.006	0.025 ± 0.004
		N	42	42	21	61	42
	Flaps (1 µm)	min–max	1.07–1.78	1.17–3.36	0.88–2.18	0.32–1.50	2.11–3.55
		mean ± SD	1.46 ± 0.20	1.78 ± 0.57	1.66 ± 0.51	1.03 ± 0.51	2.76 ± 0.48

		N	14	13	6	8	16
	Subcentral pore diameter (µm)	min–max			0.056–0.13	0.09–0.127	
		mean ± SD	absent	absent	0.089 ± 0.03	0.113 ± 0.01	absent
		N			13	13	
Annulus	**Dia (µm)**	min–max	0.11–0.46	0.20–0.57	0.38–0.54	0.50–1.20	0.11–0.26
		mean ± SD	0.29 ± 0.13	0.40 ± 0.12	0.45 ± 0.07	0.77 ± 0.24	0.19 ± 0.06
		N	11	12	7	15	13
	Total areolae number	min–max	5–25	12–18	22–40	14–50	8–16
		mean ± SD	13.83 ± 7.55	14.67 ± 1.67	30.0 ± 5.92	34.20 ± 11.71	10.62 ± 2.33
		N	12	12	7	15	13
	Central areolae diameter (µm)	min–max	0.02–0.04	0.020–0.035	0.03–0.04	0.03–0.055	0.02–0.03
		mean ± SD	0.03 ± 0.002	0.029 ± 0.003	0.035 ± 0.003	0.07 ± 0.16	0.025 ± 0.003
		N	42	42	21	33	42
Valve mantle	**Height (µm)**	min–max	0.5–1.07	1.1–1.66	0.38–0.90	0.27–1.80	0.23–0.4
		mean ± SD	0.85 ± 0.25	1.42 ± 0.19	0.68 ± 0.21	1.14 ± 0.55	0.316 ± 0.048
		N	13	15	7	16	42
	Striae (rows in 1 µm)	min–max	8–10	8–11	10–13	7–10	
		mean ± SD	9.69 ± 0.63	8.80 ± 0.94	10.86 ± 1.21	8.31 ± 0.87	absent
		N	13	15	7	16	
	Radial areolae (Total)	min–max	7–14	14–21	12–20	9–20	
		mean ± SD	10.77 ± 1.54	16.53 ± 2.29	15.00 ± 2.76	14.94 ± 3.60	scarse
		N	13	15	6	16	
	Areolae diameter (µm)	min–max	0.03-0.045	0.03-0.045	0.036–0.04	0.035-0.04	
		mean ± SD	0.04 ± 0.014	0.036 ± 0.004	0.038 ± 0.001	0.04 ± 0.002	-
		N	42	42	21	29	
Copulae	**Length (µm)**	min–max	7.6–15.9	8.41–20.5	6.66–17.7	11–28.8	3.61–5.58

	mean ± SD	10.40 ± 3.23	12.28 ± 2.28	13.66 ± 4.13	20.45 ± 5.08	4.39 ± 0.69
	N	5	19	5	17	10
Width (µm)	min–max	1.2–2.1	0.90–3.65	1.2–3.39	2.5–6.5	1.51–2.34
	mean ± SD	1.62 ± 0.37	2.07 ± 0.78	2.10 ± 0.78	3.70 ± 0.88	1.82 ± 0.30
	N	5	20	10	17	10
Striae (rows in 1 µm)	min–max	11–16	9–15	10–15	8–12	14–15
	mean ± SD	13.60 ± 1.95	11.70 ± 1.69	12.50 ± 1.72	10.71 ± 0.99	14.60 ± 0.52
	N	5	20	10	17	10
Areloae number (1 in µm)	min–max	11–13	10–16	11–18	11–14	15–17
	mean ± SD	12.8 ± 1.10	12.70 ± 1.53	14.10 ± 2.60	11.88 ± 0.93	15.60 ± 1.17
	N	5	20	10	17	10
	min–max	0.03–0.045	0.0215–0.035	0.03–0.05	0.03–0.04	0.032–0.045
Areolae diameter (µm)	mean ± SD	0.036 ± 0.004	0.032 ± 0.001	0.031 ± 0.09	0.037 ± 0.002	0.04 ± 0.004
	N	42	41	31	40	42

All information is from culture material

mantle, where the areolae become less dense and smaller (Fig. 2.1, E and G; Table 2.5). The valve surface is smooth except for a ring of spines or flap-like, triangular structures at the junction of the valve face with the mantle (Fig. 2.1, D–G). The valve mantle is smooth and has parallel striae (8–10 μm^{-1}) of round to rectangular areolae (Fig. 2.1, F and G). The mantle evenly curves from the valve face to join the valvocopula (Fig. 2.1, E–G). The cell girdle is made of thin, nearly trapezoidal half bands. A thin hyaline ridge runs parallel and close to the margins of the three shorter sides of the band, delimiting one or a few parallel rows of areolae (Fig. 2.1, H). At times the ridge is only visible along the two oblique sides of the band. Crosswise oriented striae are present on the bands at a density of 11–16 in 1 μm, becoming less regular and dense towards the ridge (Fig. 2.1, H). Vegetative cell size expansion through auxospore-like structures occurred in cultures growing at constant rate. No spore formation was observed under nutrient depletion conditions.

2.3.2. ***Leptocylindrus convexus*** Nanjappa et Zingone, sp. nov. (Fig. 2.2, A–J)

Diagnosis. Cells cylindrical, 3–8 μm in diameter, 22–65 μm in pervalvar length, forming filamentous chains or often found in couples. Plastids few, ellipsoidal-lanceolate and elongated along the pervalvar axis, often located in the central part of the cell. Valve faces slightly concave or convex. Annulus occasionally seen as an irregular hyaline ring delimiting a number of areolae in the centre of the valve face. No distinct sub-central pore. Poroid areolae (10–14 in 1 μm) on the valve face in radiating striae (10–14 in 1 μm). Mantle wide, slanting abruptly from the valve face margin towards the valvocopula, with areolae in parallel rows (8–11 in 1 μm). Ring of short, triangular and often blunt spines at the junction of valve face and mantle. Girdle bands nearly trapezoidal half-bands, with irregular crosswise rows of areolae. Resting spores not observed.

Holotype: A permanent slide of strain SZN-B781 deposited at the SZN Museum as no. SZN-B781-01.

Iconotype: Fig. 2.2, A–J

Materials examined: Strains SZN-B781 and SZN-B772.

Type locality: GoN, Italy (South Tyrrhenian Sea, Mediterranean Sea).

Etymology: The species epithet *convexus* (convex) refers to the shape of the valve.

Description. The cells are 3–8 μm in diameter, 22–65 μm in pervalvar length, and join by the valve face to form short filaments generally of 2 or 3 cells (Fig. 2.2, A–C and Table 2.4). Chains of up to 68 cells were occasionally observed, at times in gently undulated or rarely spiralling chains. Each cell possesses few (3–11), linear or elliptical-lanceolate plastids (Fig. 2.2, A–C) up to 10–15 μm long in thin (3–4 μm diameter) cells. The plastids are often arranged in a star-like pattern around the central nucleus (Fig. 2.2, C). Valves are either convex or concave, with a markedly thicker, electron-dense central part and an irregular, hardly visible central annulus (Fig. 2.2, E and G). Striae of areolae radiate from the central part of the valve and generally continue across the mantle, where they run in parallel rows of round to rectangular areolae (Fig. 2.2, E and G, and Table 2.5). Spines or flap like, triangular structures are seen at the margin of the valve face, but they are generally blunter and denser than in other species (Fig. 2.2, F). The valve mantle is wide and slanting towards the valvacopula. This gives a markedly convex outline to the whole valve (Fig. 2.2, D and F), resulting in a pronounced constriction at the junction between sibling cells, which is also discernible in the light microscope (Fig. 2.2, C). Copulae are made of half bands, nearly trapezoidal (Fig. 2.2, I and J) or rarely ribbon-like (Fig. 2.2, H), whose width at times decreases from the valve towards the mid of the girdle. A thin hyaline ridge runs parallel to the shorter sides of the band, at times only seen along the two oblique ones, delimiting one or a few parallel rows of areolae (Fig. 2.2, H–J, Table 2.5). Crosswise oriented striae are present on the bands at a density of 9–15 in 1 μm, becoming less regular and dense towards the ridge

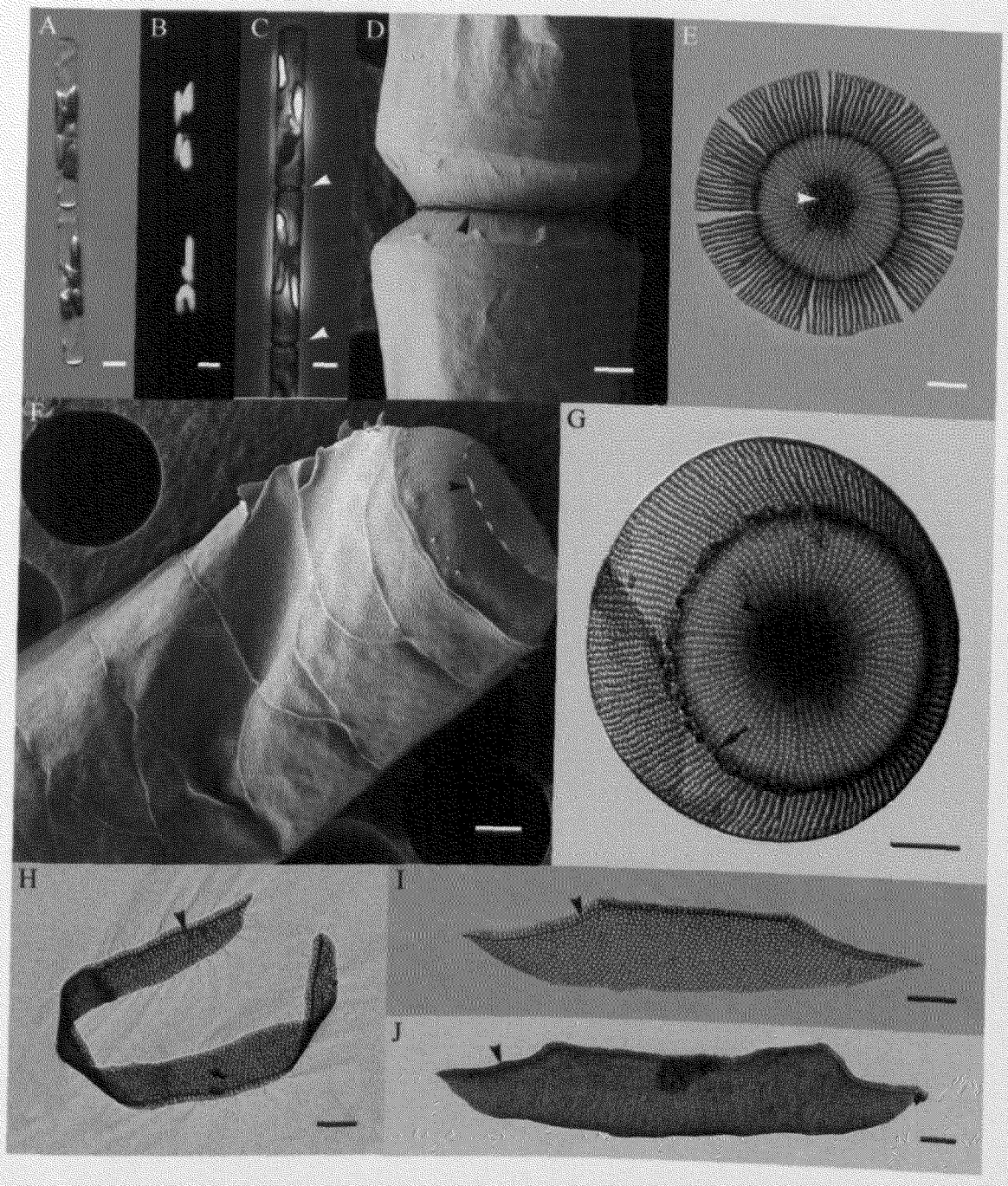


Fig. 2.2. *Leptocylindrus convexus,* strain SZN-B768. LM, DIC: A; LM, epifluorescence: B; LM, phase contrast: C. SEM: D and F. TEM: E, G–I and J. (A) Colony of two cells. Scale bar, 5 µm. (B) Same colony as in A; note the shape of the plastids. Scale bar, 5 µm. (C) Fragment of a colony; note the marked constriction at the cell joinings (arrowheads). Scale bar, 5 µm. (D) Details of two sibling cells joined by the valve face; note the wide mantle and the flaps in alternated position at the border of the valve faces (arrowhead). Scale bar, 1 µm. (E) Valve with mantle breaks due to the marked convexity; note the absence of subcentral pore and the presence of a faint annulus (arrowhead). Scale bar, 1 µm. (F) Detail of a cell in girdle view, with flaps (arrowhead) around the valve face, very convex mantle and broad, pointed copulae. Scale bar, 1 µm. (G) Valve with an almost continuous ring of flaps (arrow) and an electron dense

central area (arrowhead). Scale bar, 1 μm. (H, I and J) Nearly trapezoidal copulae, with different size and shape; note the almost continuous hyaline line close to the borders (arrowheads). Scale bar, 1 μm.

(Fig. 2.2, J). Resting spore formation was not observed in nutrient depleted media. Cell enlargement was not directly observed either, although size did vary over the time in individual cultures.

2.3.3. *Leptocylindrus danicus* Cleve 1889 (Fig. 2.3, A–I and 2.4 A–F)

Emended diagnosis. Cells cylindrical, 4.5–7.4 μm in diameter, forming filamentous colonies. Plastids numerous, lens-shaped or ellipsoidal. Valves with convex or concave faces and a central annulus delimiting a group of poroid areolae. Distinct sub-central pore adjacent to the annulus. Poroid areolae (18–30 in 1 μm) arranged in radiating striae (11–16 in 1 μm). Mantle evenly curving proximally and continuing almost perpendicular to the valve face distally. Ring of short, triangular and often blunt spines at the junction of valve face and mantle. Girdle comprised of elongated, nearly trapezoidal half bands, with areolae along irregular crosswise striae. Auxospore nearly spherical, smooth, covered with weakly silicified circular plates. Spores semi-globular, composed of two unequal valves bearing triangular or pyramidal spines, with smooth or at times serrated margins.

Neotype: A permanent slide of strain SZN-B650 is deposited in the SZN Museum as no. SZN-B650-01.

Iconotype: Figure 2.3, A–I

Materials examined: Strains SZN-B650, SZN-B707 and SZN-B714.

Type locality: Kattegat Bay.

Description. The cells are 3–13 μm in diameter and 22–75 μm in pervalvar length (Table 2.4), and form filaments often composed of hundreds (<165, Table 2.4) cells. Each cell possesses many (7–36) discoid, rarely ovoid plastids (Fig. 2.3, A and B), which are regularly distributed across the pervalvar axis along the cell periphery. The

valve faces are convex or concave and have a slightly more silicified central area, with a group of areolae delimited by a hyaline annulus (Fig. 2.3, F and G). On the valve, the striae originate from the central annulus and run across the mantle, at times splitting into two half-way on the valve face (Fig. 2.3, F). The sub-central pore is conspicuous, 0.06–0.13 μm, with a hyaline margin which is also visible on the internal surface of the valve (Fig. 2.3, G) and is generally located adjacent to the hyaline ring (at least in 80% of the observed valves). A circle of spines or flap-like, triangular structures are found at the margin of the valve face (Fig. 2.3, C–E). The mantle bends proximally, and then runs almost perpendicular to the valve face towards the valvocopula (Fig. 2.3, C). It is perforated by parallel striae of round areolae, which almost always are in continuity with those of the valve face (Fig. 2.3, D and F). The girdle consists of nearly trapezoidal half-band segments, with irregular crosswise striae of poroid areolae. A thin hyaline ridge runs parallel to the margins of the two oblique sides of the copulae (Fig. 2.3, H and I), while one or more irregular ridges are seen at times along the shorter of the two parallel sides. Crosswise or obliquely oriented striae (10–15 in 1 μm) are present on the bands, becoming less regular towards the ridge (Fig. 2.3, H and I). Under nutrient-depleted conditions, spherical auxospores covered by siliceous scales (Fig. 2.4, A) are produced which turn into semi-globular spiny spores (Fig. 2.4, B and C). Each spore is composed of two unequal valves, a bigger semi-circular epivalve (Fig. 2.4, C and E) and a relatively smaller hypovalve (Fig. 2.4, D). Both the epivalve and hypovalve have triangular or pyramidal spines which have smooth or at times branched or serrated margins (Fig. 2.4, D–F).

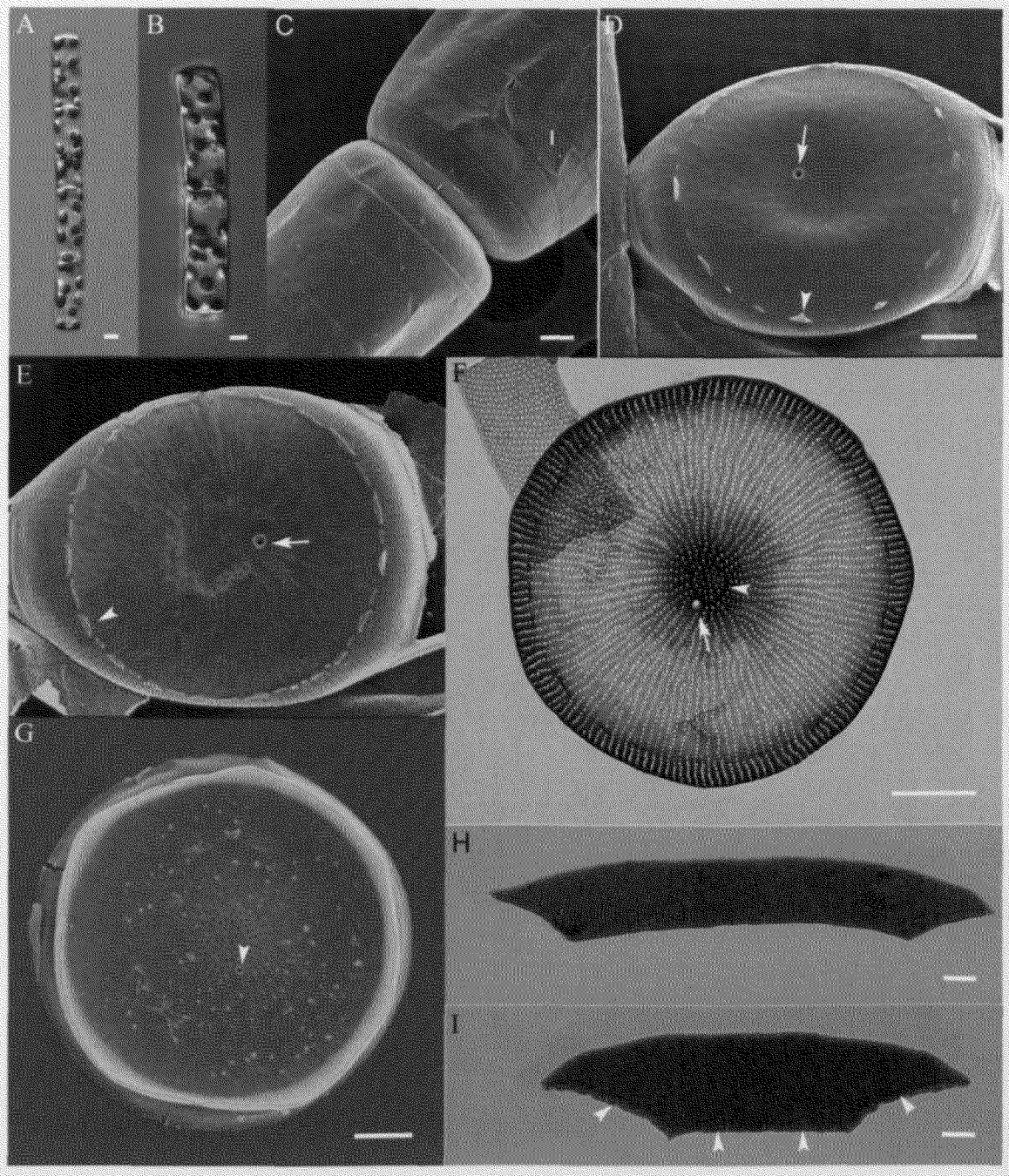


Fig. 2.3. *Leptocylindrus danicus,* strain SZN-B650. LM, DIC: A and B. SEM: C–E and G. TEM: F, H and I. (A) Colony of four cells. Scale bar, 5 µm. (B) Colony of two cells showing the convex or concave valve surface and discoidal plastids. Scale bar, 5 µm. (C) Detail of two cells joined by the valve face, with hardly visible flaps. Scale bar, 1 µm. (D) Cell in valve view, with subcentral pore (arrow) and spaced triangular flaps (arrowhead). Scale bar, 1 µm. E) Cell in valve view, with subcentral pore (arrow) and almost continuous line of flaps (arrowhead). Scale bar, 1 µm. (F) Valve with annulus (arrow) and adjacent sub-central pore (arrowhead). Scale bar, 1 µm. (G) Internal view of a valve with the sub-central pore (arrowhead). Scale bar 1 µm. (H and I) Nearly trapezoidal copulae, with different length and shape; note the almost continuous hyaline line close to one border (arrowheads) Scale bar, 1 µm.

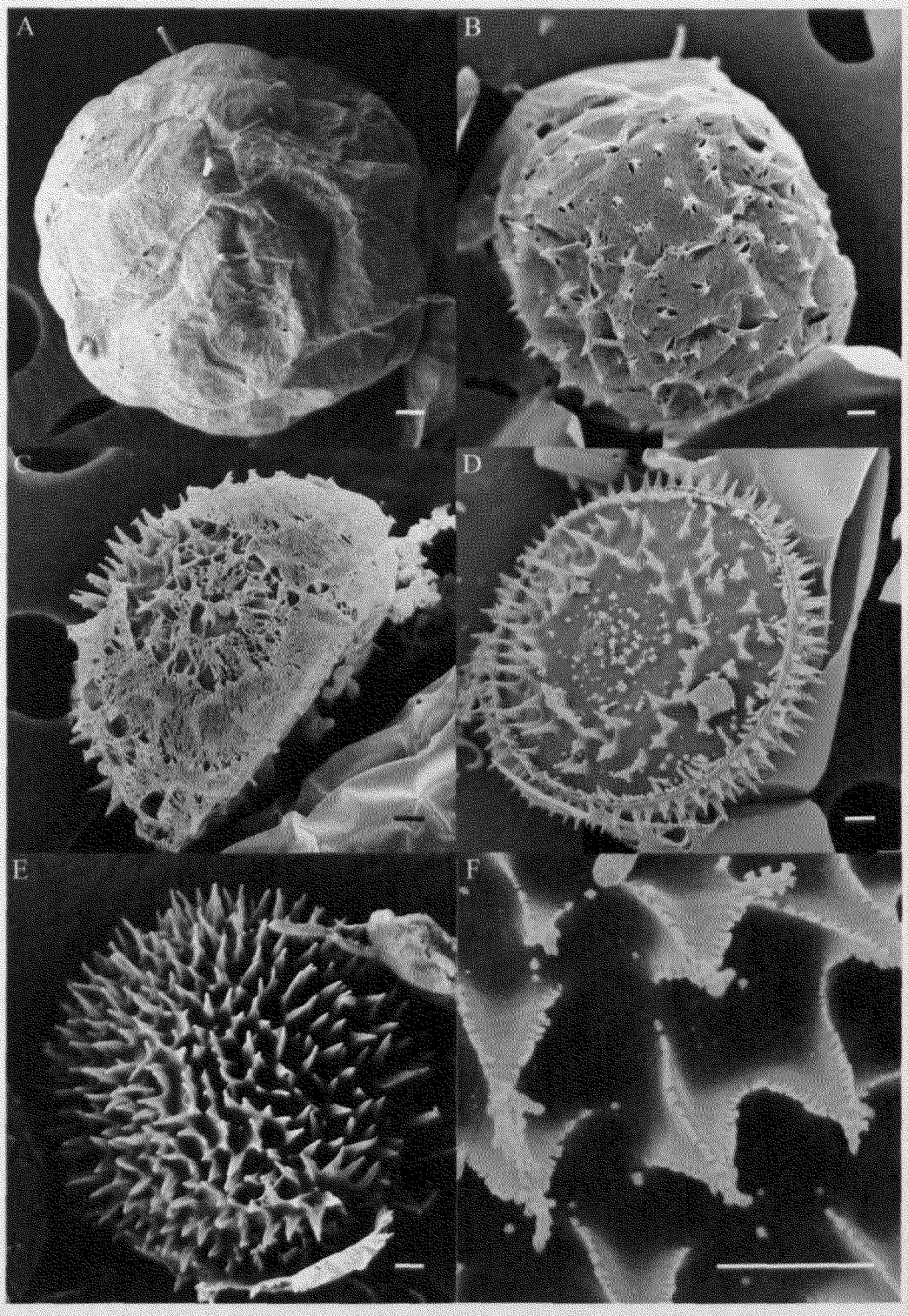


Fig. 2.4. Spore morphology of *L. danicus*, strain SZN-B650. SEM. (A) Auxospore covered with finely areolated silica scales. Scale bar, 1 µm. (B) Spiny spore forming inside the auxospore. Scale bar, 1 µm. (C) Spore in girdle view, with fragments of scales still attached to the spines; note the semi-globular shape. Scale bar, 1 µm. (D) Hypovalvar view of a spore. Scale bar, 1 µm. (E) Epivalvar view of a spore. Scale bar, 1 µm. (D) Detail of a group of pyramidal epivalvar spines, with serrated margins. Scale bar, 1 µm.

2.3.4. *Leptocylindrus hargravesii* Nanjappa et Zingone sp. nov. (Fig. 2.5, A–K and 2.6, A–D)

Diagnosis. Cells 3–15 μm in diameter, with heavily silicified, cylindrical frustules, joined to form filamentous chains. Plastids numerous, lens-shaped to ellipsoidal. Valves with convex or concave faces and a central annulus delimiting a group of poroids. Distinct sub-central pore generally not adjacent to the annulus. Poroid areolae (10–14 in 1 μm) arranged in radiating striae (6–11 in 1 μm), in most cases interrupted before reaching the mantle. Mantle evenly curved proximally, almost perpendicular to the valve face distally. Ring of short, triangular and often blunt spines at the junction of valve face and mantle. Girdle comprised of elongated, often nearly trapezoidal half-bands with areolae along irregular crosswise rows. Auxospore nearly spherical, smooth, covered with weakly silicified circular plates. Spores semi-globular, composed of two unequal valves. Spore valves with pyramidal spines with smooth or at times serrated margins.

Holotype: A permanent slide of strain SZN-B781 has been deposited in the SZN Museum as no. SZN-B781-01.

Iconotype: Figure 2.5, A–K and 2.6, A–D.

Materials examined: Strains SZN-B781, SZN-B772 and SZN-B773.

Type locality: GoN, Italy (South Tyrrhenian Sea, Mediterranean Sea).

Etymology: The species is dedicated to Paul Hargraves, who greatly contributed to the understanding of the diversity and life cycle of the genus *Leptocylindrus*.

Description. The cells are 3–15 μm in diameter and 30–90 μm in pervalvar length (Table 2.4), with a relatively stout frustule (Fig. 2.5, A–C). In culture, filamentous chains may be composed of up to 162 cells (Table 2.4). Each cell possesses numerous (9–55, Table 2.4), discoid or ellipsoidal plastids, regularly distributed along the cell periphery (Fig. 2.5, A and B). Valves have concave or convex valve face. The areolae

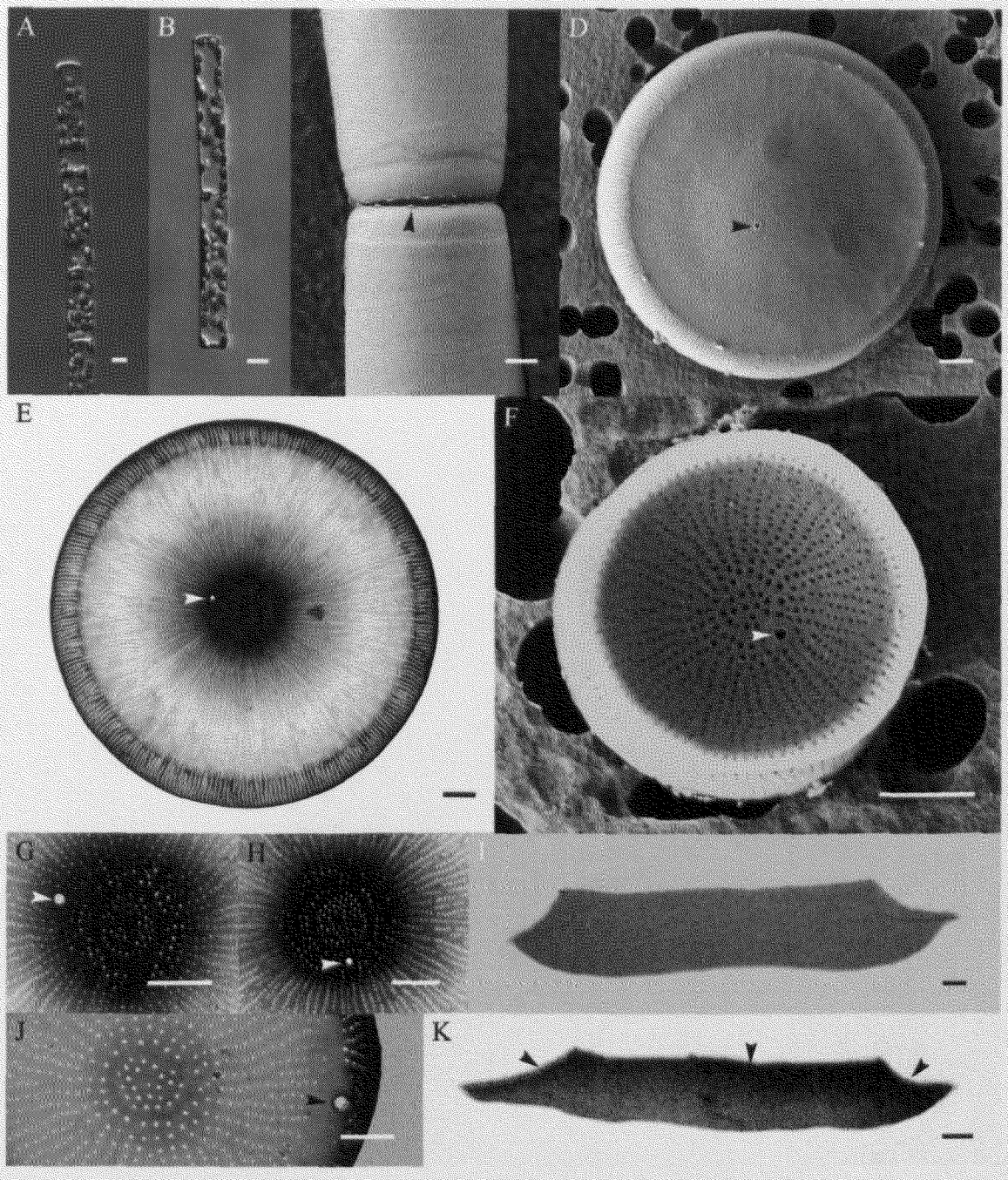


Fig. 2.5. *Leptocylindrus hargravesii,* strain SZN-B781. LM, DIC: A and B. SEM: C, D and F. TEM: E and G–K. (A) Colonial cells. Scale bar, 5µm. (B) Colony of two cells showing the convex or concave valve surface and discoidal plastids. Scale bar, 5µm. (C) Detail of two cells joined by the valve face, with flaps alternating along the marginal ring (arrowhead). Scale bar, 1 µm. (D) Valve with subcentral pore and hardly visible flaps. Scale bar, 1 µm. (E) Valve with sub-central pore (arrowhead); note the heavily silicified central area. Scale bar, 1µm. (F) Internal view of a valve, with the subcentral pore (arrowhead). Scale bar, 1µm. (G, H, and J) Details of valves, showing the variable position of the subcentral pore. Scale bar, 1 µm. (I and K) Heavily silicified, nearly trapezoidal copulae of different length and shape; note the almost continuous hyaline line close to one border (arrowheads). Scale bar, 1 µm.

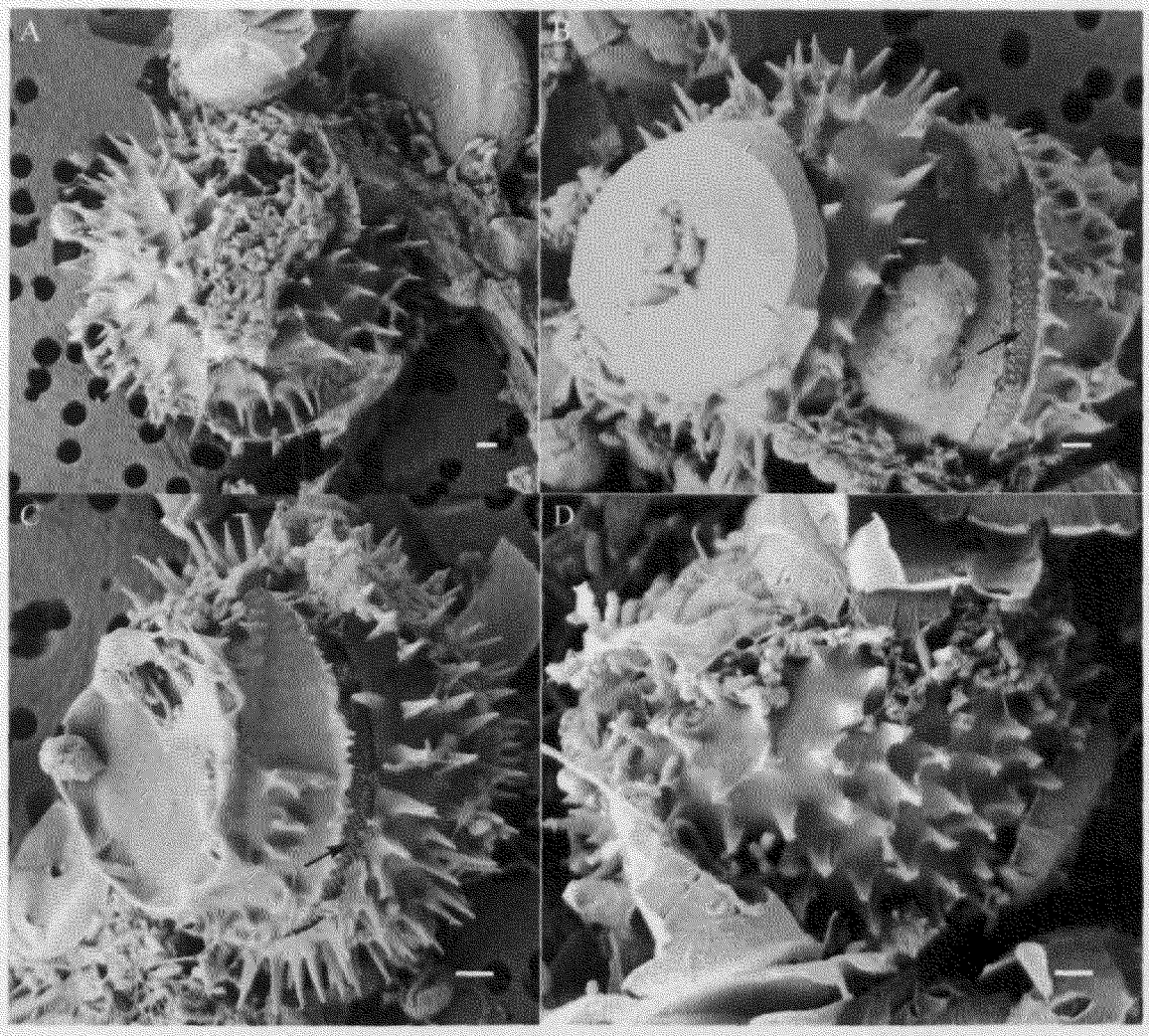


Fig. 2.6. Spore morphology of *L. hargravesii*, strain SZN-B781. SEM. (A) Spiny spore with scale fragments. Scale bar, 1 μm. (B) From the right, a cell valve, a spore epivalve and a whole spore in girdle view; note the granules on the hypovalvar rim (arrow). Scale bar, 1 μm. (C) Spore in oblique girdle view, with a rather smooth hypovalve; note the granulated hypovalvar rim (arrow). Scale bar, 1 μm. (D) Epivalve with fragments of valves and scales attached to the spines. Scale bar, 1 μm.

on the valve face (10–12.5 in 1 μm) are simple round pores placed in uniseriate radiating striae (6–11 in 1 μm) which are often interrupted distally, generally not reaching the mantle. The density and size of the areolae on the striae also tend to diminish towards the mantle (Table 2.5). A thicker hyaline annulus delimiting a group of areolae (Fig. 2.5, G, H and J) is present in the mid of the valve face, which is at times hardly visible since these area can be heavily silicified (Fig. 2.5, E). A prominent pore (0.09–0.13 μm) is observed close to the annulus but never adjacent to it. It has a thickened margin which is also seen on the internal surface of the valve (Fig. 2.5, F). A

variable position in culture material is at times observed for this pore, which can rarely be found close to the valve face margin (Fig. 2.5, G, H and J). Marginal spines or flap like, triangular structures are present along the border between mantle and valve face (Fig. 2.5, C). The mantle is set off from valve face at an angle of about 90°, proximally bending towards the valve face (Fig. 2.5, C and D). It has round to rectangular, areolae arranged in parallel striae, which often are not continuous with those of the valve face. The girdle is composed of numerous pointed, nearly trapezoidal bands. A thin hyaline ridge runs parallel to one margin of the band, generally opposite to the longer side, delimiting one or a few parallel rows of areolae (Fig 2.5, I and K). A ridge at times can be seen only along the oblique sides of the band. Crosswise oriented striae are present on the bands at a density of 8–12 in 1 μm, becoming less regular towards the ridge (Fig. 2.5, I).

Spiny, semi globular spores are formed upon induction by simple deprivation of nutrients (Fig. 2.6, A–I). These spores develop inside smooth auxospores, which show disc-like scales. Each resting spore consists of two unequal sized valves, a lager epivalve and a smaller hypovalve (Fig. 2.6, B). Both valves are heavily silicified and covered with pyramidal spines with smooth or at times serrated margins and residual scale fragments sticking to them. The hypovalve can be smoother than the epivalve, and has a marginal band covered with granules arranged in regular rows, which fits into the epivalve (Fig. 2.6, B and C).

2.3.5. *Tenuicylindrus* Nanjappa et Zingone gen. nov.

Tenuicylindrus belgicus Nanjappa et Zingone, comb. nov. (Fig. 6, A–J)

Synonym: ***Leptocylindrus belgicus*** Meunier 1915

Diagnosis: Cylindrical cells, 2–2.5 μm in diameter, forming short chains, slightly undulating in culture material but always straight in nature. Two thin, elongated, leaf-shaped plastids per cells. Two granules visible at each end of the cells, under the valve

face. Circular valves, with radial striae (7–11 in 1 μm) of pores. Mantle perpendicular to the valve face, smooth and with scarce pores, proximally ending with a zig-zag margin joining the valve face rim. Triangular teeth-shaped processes at the boundary between mantle and valve faces, closely fitting those of the sister cells. Girdle composed of lip-shaped half bands with longitudinal rows of pores. Spores not observed.

Holotype : Plate XII, Fig. 4 in Meunier 1915

Epitype: A permanent slide of strain SZN-b has been deposited in the SZB Museum as no. SZN-B739.

Iconotype: Figure 2.7, A–J

Materials examined: Strains SZN-B739, SZN-B755 and SZN-B756.

Type locality: Belgian waters.

Etymology: The genus name refers to the extremely thin (*tenuis* in Latin) frustules of this species. The species epithet assigned by Meunier (1915) refers to the type locality of the species.

Description. Cells are cylindrical, with small and scarcely variable diameter (2–2.5 μm) and relatively long pervalvar axis (23–50 μm, Table 2.4). Each cell has two long and narrow leaf-like plastids, positioned on either sides of the nucleus along the girdle (Fig. 2.7, A and B). Cells are joined by the valves to form long filaments, which in natural samples are always straight and comprised of two or three cells, whereas in culture they are often slightly curved or undulated and are comprised of many cells (2–14 cells per colony, Table 2.4). The frustule is delicate and only requires a weak acid treatment to eliminate the organic matter and preserve the morphological features. The valve face is either convex or concave (Fig. 2.7, C and D), the concave valves fitting the convex ones of the sister cell in the chains. The central area of the valve has a group of areolae that at times is ill defined or may be delimited by a hyaline, poreless ring. Over the valve face, round areolae of variable size are placed along radiating striae (Fig. 2.7, F). The mantle

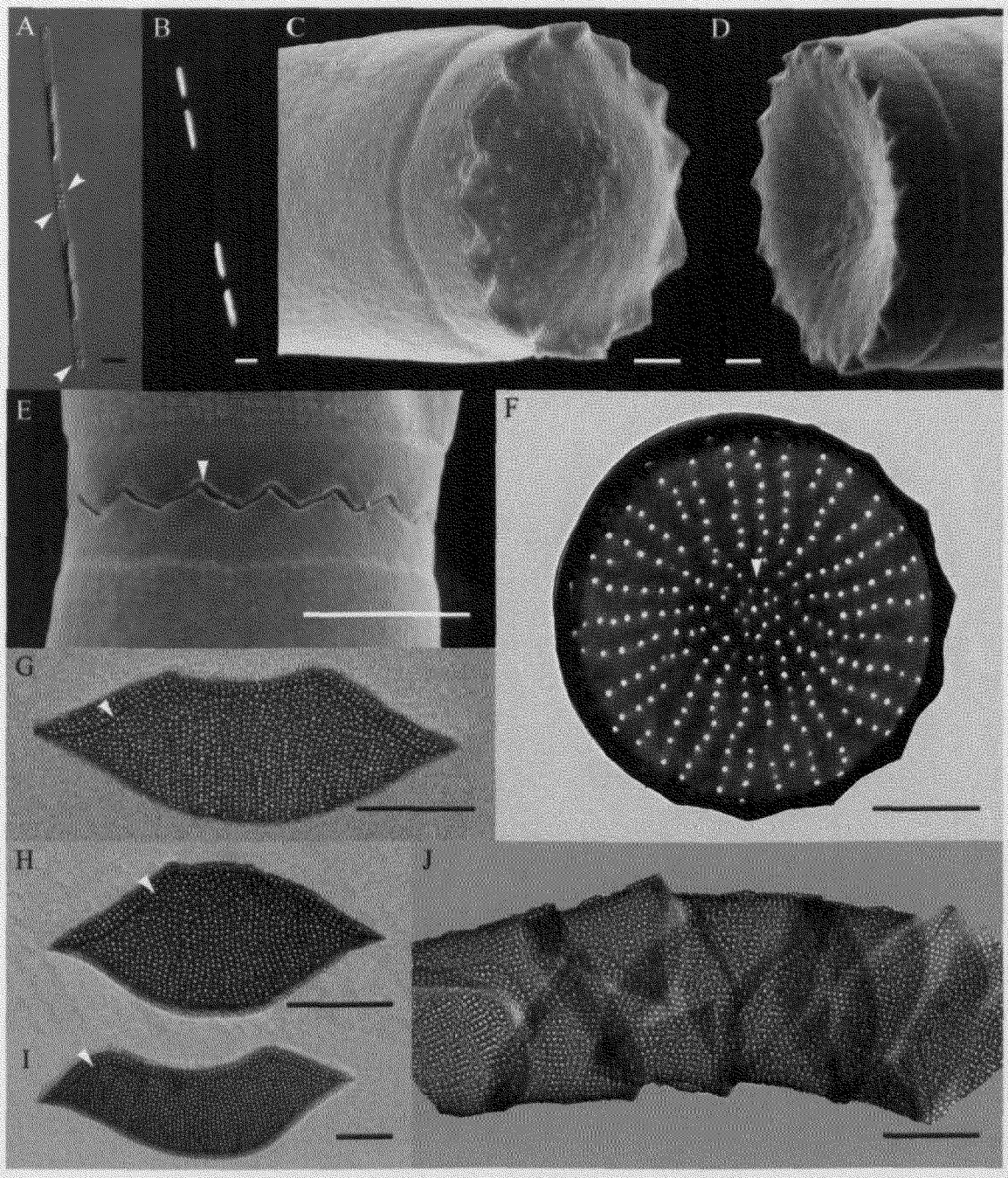


Fig. 2.7. *Tenuicylindrus belgicus*, strain SZN-B739. LM, DIC: A; epifluorescence: B. SEM: C, D and E. TEM: F–J. (A) Colony of two cells; note the very elongated plastids and a couple of granules at both the cell ends (arrowheads). Scale bar, 5 µm. (B) Same colony as in A; note the shape of the plastids. Scale bar, 5 µm. (C) Valve with convex valve face and the crown-shaped mantle. Scale bar, 0.2 µm. (D) Valve with concave valve face. Scale bar, 0.2 µm. (E) Detail of the junction between two sibling cells; note the zig-saw fit (arrowhead) and the smooth, poreless mantle. Scale bar, 1 µm. (F) Valve ultrastructure. Scale bar, 1 µm. (G) Valvacopula; note the hyaline line close to the border (arrowhead). Scale bar, 1 µm. (H) Copula; note the hyaline line close to the border (arrowhead). Scale bar, 1 µm. (I) Copula with a less common shape; note the hyaline line close to the border (arrowhead). Scale bar, 1 µm. (J) Cell in girdle view; note the arrangement of the copulae. Scale bar, 1 µm.

is smooth, with very sparse areolae, and lies perpendicular to the valve face plane. The margin towards the valvacopula is smooth, while a sharp, zigzag-edged margin is present towards the valve face, conferring the valve a crown-like shape. The zigzag margin of one cell perfectly fits that of the adjacent cell, thus producing chains with no constrictions at the boundary between two sister cells (Fig. 2.7, E). The girdle bands have pointed ends and often show a lip-like outline. They bear longitudinal, more or less regular, striae of areolae. One edge of the band is smooth while the other edge is finely serrated (Fig. 2.7, H, I and K). Cultured cells maintain a rather constant cell size range, with no indication of vegetative autoenlargement or of auxospore formation. Spores were not observed under nutrient-deprived conditions.

2.3.6. Seasonal distribution

Based on the time the 85 strains were brought in cultivation and on their genetic identification (Table 2.1), the seasonality for the five species in the GoN was reconstructed. This was possible considering that isolation of *Leptocylindrus* and *Tenuicylindrus* strains was regularly performed over more than one year, and was also accompanied by careful observations of the net samples on a weekly basis. For some weeks of the year *Leptocylindrus* spp. cells in net samples were scarce and culturing was unsuccessful. Microscopic observations on those rare cells often did not allow species identification thus generating gaps in the seasonal distribution map. *Leptocylindrus aporus* strains were all retrieved from mid-July to mid-November, whilst *L. convexus* strains were found from January to towards the end of March. *Leptocylindrus danicus* showed the longest period of occurrence across seasons, with strains retrieved from mid-November to mid-July, whilst *L. hargravesii* were only retrieved in December and January. Finally, *Tenuicylindrus belgicus* strains were retrieved from the end of August to the beginning of November. The period of

occurrence of the individual species is plotted in Fig. 2.8 against the temperature and day length value that are typical for the seasonal cycle at station LTER-MC.

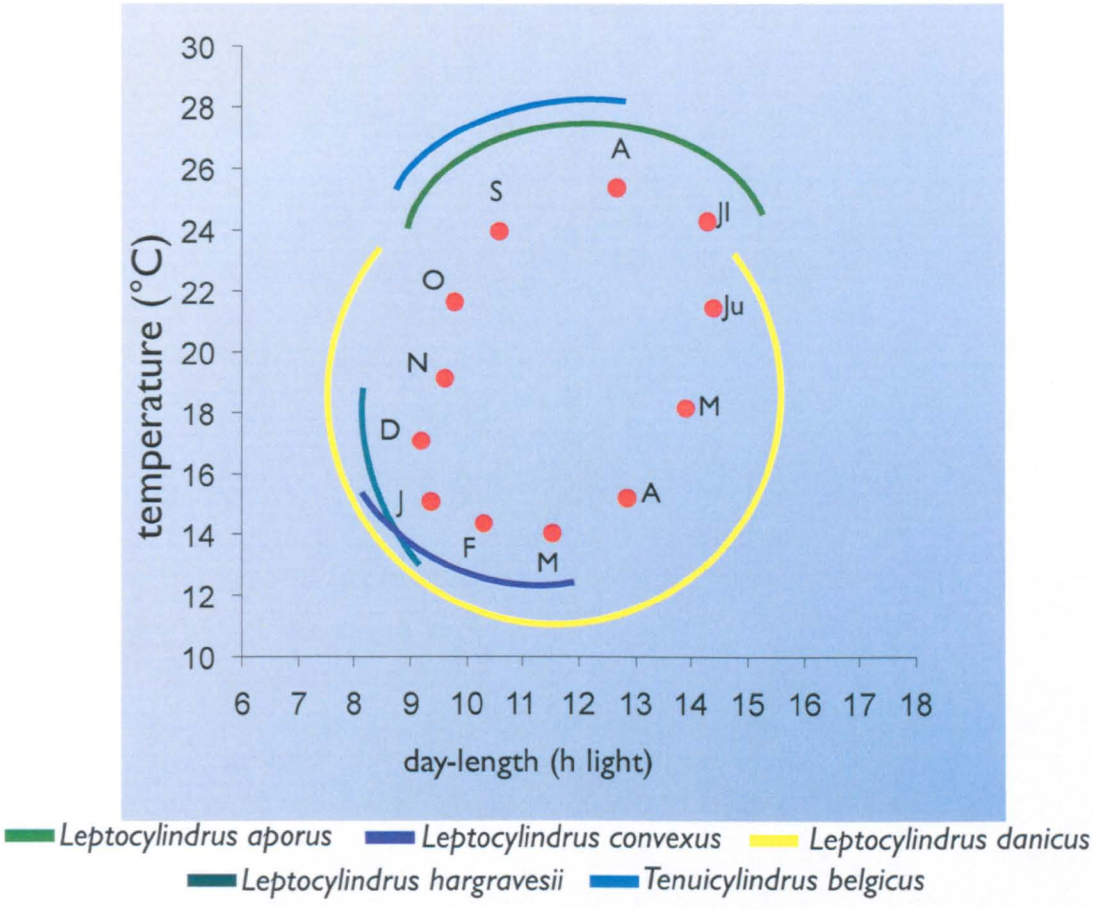


Fig. 2.8. Temporal distribution of the *Tenuicylindrus* and *Leptocylindrus* species at the station LTER-MC, in the Gulf of Naples, plotted on a temperature/photoperiod diagram.

2.3.7. *Molecular phylogenies.* Information about the aligned sequences of the nuclear encoded SSU rDNA, the partial LSU rDNA and the ITS region, as well as the plastid encoded SSU rDNA rbcL and psbC of the various *Leptocylindrus* and *Tenuicylindrus* strains have been listed in Table 2.1. The ITS1 and ITS2 sequences within the ITS region were unalignable among different species. For *L. danicus* and *L. hargravesii* only, however, the ITS sequences, were alignable and differed at least at 29 positions. ITS sequences within each of the species remained identical. The 5.8S rDNA within the ITS region aligned well across all the *Leptocylindrus* sequences, but also this conserved region differed among the species.

Overviews of the phylogenetic relationships among the *Leptocylindrus* and *Tenuicylindrus* species inferred from the six different molecular markers are shown in Fig. 2.9 (as cladograms) and Fig. 2.10 (as phylograms). These trees resulted from analyses using only the sequences of *Leptocylindrus* and *Tenuicylindrus* and the latter have been designed as the outgroup. The number of parsimonious informative characters associated with internodes and end-branches has been represented in the cladograms in Fig. 2.9, bootstrap support for clades is presented in the phylograms in Fig. 2.10. Information about the alignment lengths and numbers of parsimonious informative positions is presented in Table 2.6. Although the number of positions varied among the different regions, almost all the regions contained sufficient information to distinguish the studied species with plastid SSU rRNA region providing the least information (57 positions). The summary of base substitutions inferred through the best fit model for different regions has been presented in the Table 2.7. All the trees showed essentially the same phylogenetic relationships, with *L. hargravesii* as sister to *L. danicus* and *L. aporus* as sister to *L. convexus*.

The following trees were inferred with sequences of *Leptocylindrus*, *Tenuicylindrus* and other genera. Selected sequences among the latter were designed as outgroup. A neighbour joining (NJ) tree showing the position of the two genera among numerous genera throughout the diatom diversity, other autotrophic Stramenopiles and heterotrophic Stramenopiles is shown in Fig. 2.11. The maximum likelihood (ML) tree inferred from the nuclear SSU rDNA sequences is shown in Fig. 2.12, the nuclear partial LSU ML-tree is shown in Fig. 2.13, the plastid SSU rDNA ML-tree is shown in Fig. 2.14, the psbC ML-tree is shown in Fig. 2.15, the rbcL ML-tree is shown in Fig. 2.16. In the nuclear-encoded SSU rDNA tree (Fig. 2.12), *Leptocylindrus* was sister to *Tenuicylindrus* (Lineage I). Within *Leptocylindrus*, the clade with *L. danicus* and *L. hargravesii* (Lineage II) branched off first, followed by a clade with *L. aporus*. *Leptocylindrus minimus* was recovered as sister to *L. convexus*. The latter three species

Table 2.6. DNA markers used in this study, their alignment lengths and the parsimony informative positions.

Gene region	Alignment length (bp)	Parsimony-informative positions
Nuclear SSU rDNA	1618	484
Partial nuclear LSU rDNA	679	354
5.8S rDNA	146	42
ITS	Unalignable*	
Partial plastid SSU rDNA	752	57
rbcL	1435	345
psbC	742	338

* total length of sequences obtained is circa 620 bp

Table 2.7. Base composition and estimated base substitution models as inferred with Modeltest (cit) for each of the alignments used in this study.

Gene region	Base substitution rates									
	A	C	G	A↔C	A↔G	A↔T	C↔G	C↔T	I	G
Nuclear SSU rDNA	0.2677	0.1863	0.2551	0.8212	2.5788	1.3773	0.7444	4.0728	0.3654	0.5570
Nuclear LSU rDNA	0.2320	0.1818	0.2980	0.9427	2.6720	1.8988	0.8757	5.9041	0.2110	0.6765
5.8S rDNA	0.2500	0.2500	0.2500	0.7959	7.4751	3.9730	0.3001	7.4751	0.5585	equal
ITS	0.2908	0.1961	0.2241	UA	UA	UA	UA	UA		
Plastid SSU rDNA	0.2887	0.1839	0.2702	1.0000	4.2159	0.5664	0.5664	7.4776	0.7337	0.5289
rbcL	0.2900	0.1551	0.2009	1.2564	4.9147	6.6717	1.7640	9.4251	0.4981	0.7194
psbC	0.2353	0.1051	0.2142	0.7892	5.2819	6.9855	0.6920	24.6727	0.6004	1.4849

UA – unalignable data

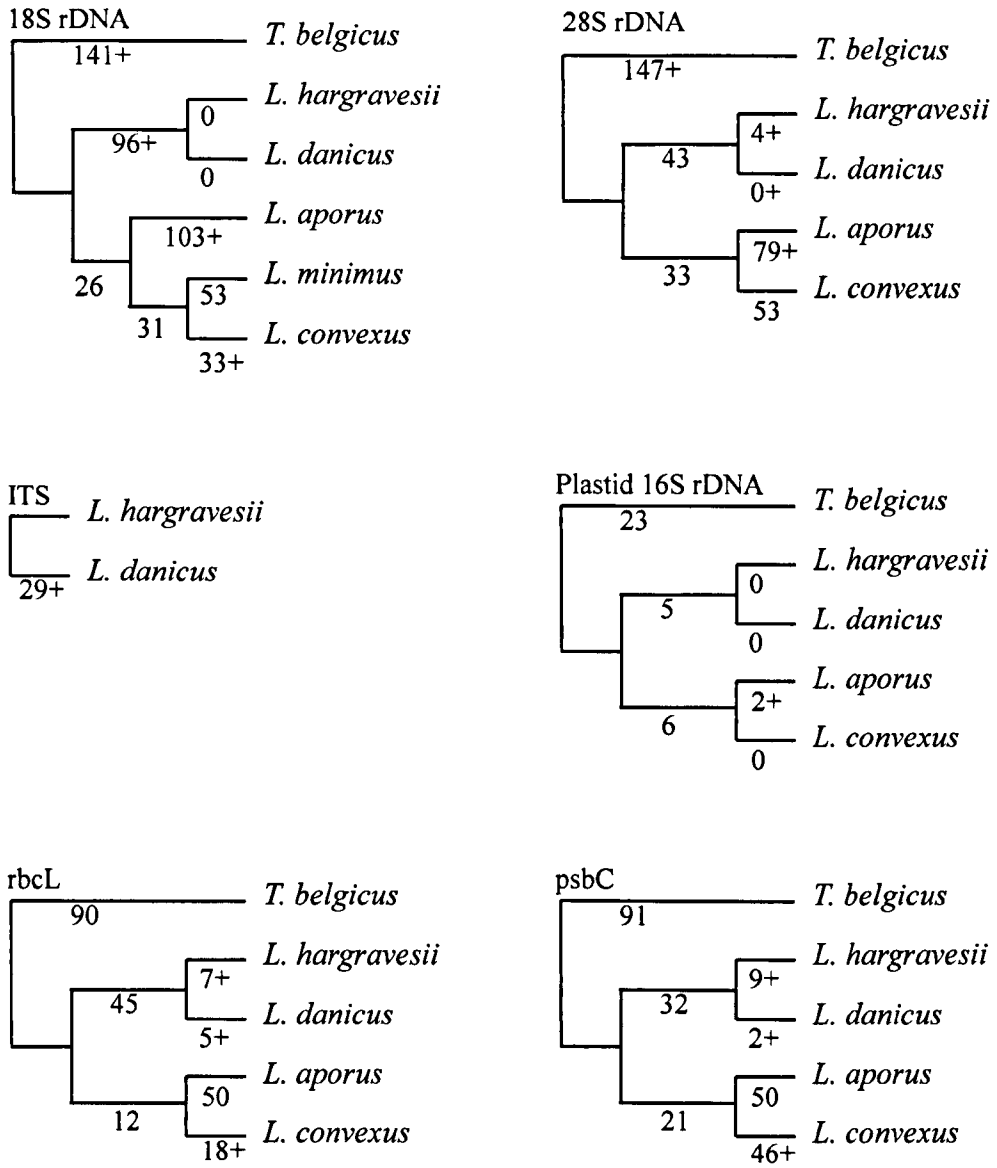


Fig. 2.9. Maximum parsimony relationships inferred for each of the alignments used in this study for *Tenuicylindrus* and *Leptocylindrus* species. Values indicate the number of parsimonious informative characters.

are designated from here onwards Lineage III. All clades within *Leptocylindrus* obtained high bootstrap support. The more extensive nuclear SSU tree in Fig. 2.11 showed the same relationships. *Leptocylindrus minimus* was missing in all other trees because only nuclear SSU rDNA sequence of this species was retrieved from GenBank.

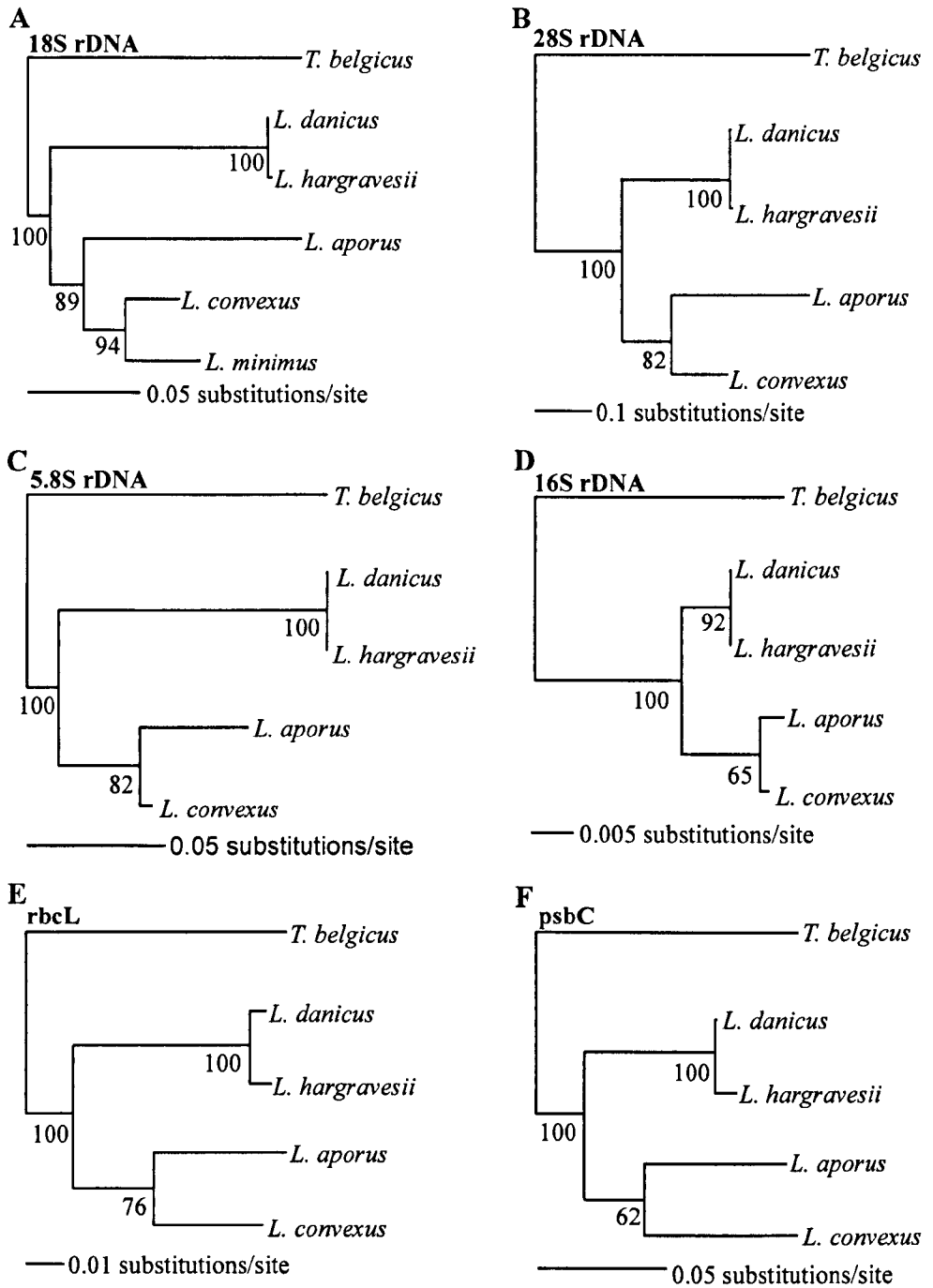


Fig. 2.10. Maximum likelihood relationships inferred for each of the alignments used in this study for *Tenuicylindrus* and *Leptocylindrus* species. Bootstrap values have been generated with 1000 replicates.

In the nuclear-encoded partial LSU-tree (Fig. 2.13), *Leptocylindrus* species grouped into a clade. *Tenuicylindrus* was not resolved as sister clade to *Leptocylindrus*, but bootstrap support for the basal clades in this tree was insufficient, thus not ruling out possible sister relationships between the two clades. All clades within *Leptocylindrus* obtained high bootstrap support.

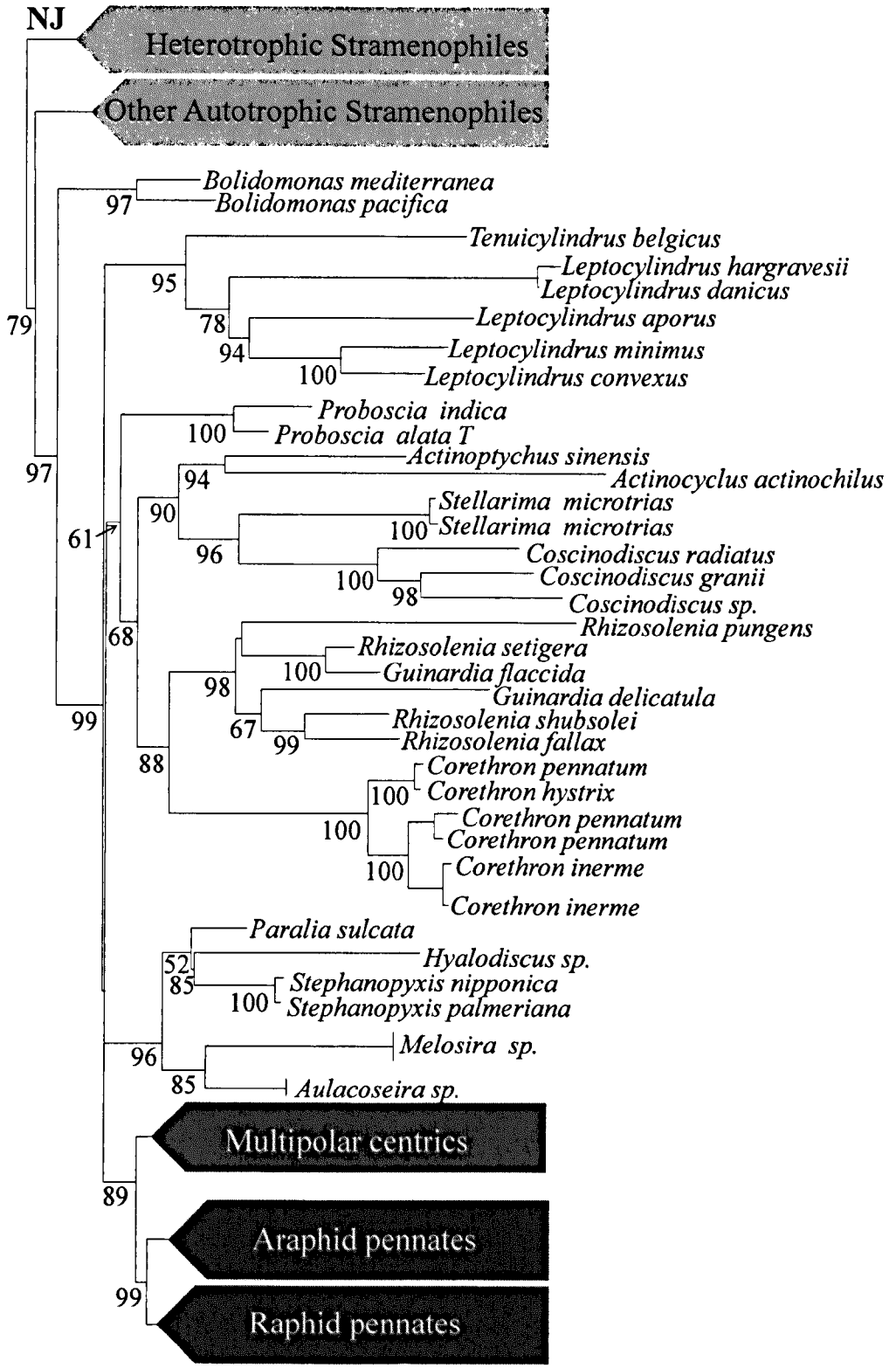


Fig. 2.11. Neighbour joining tree inferred from SSU rDNA illustrating the relationship among *Tenuicylindrus, Leptocylindrus* species and other diatom groups. Bootstrap values have been generated with 1000 replicates.

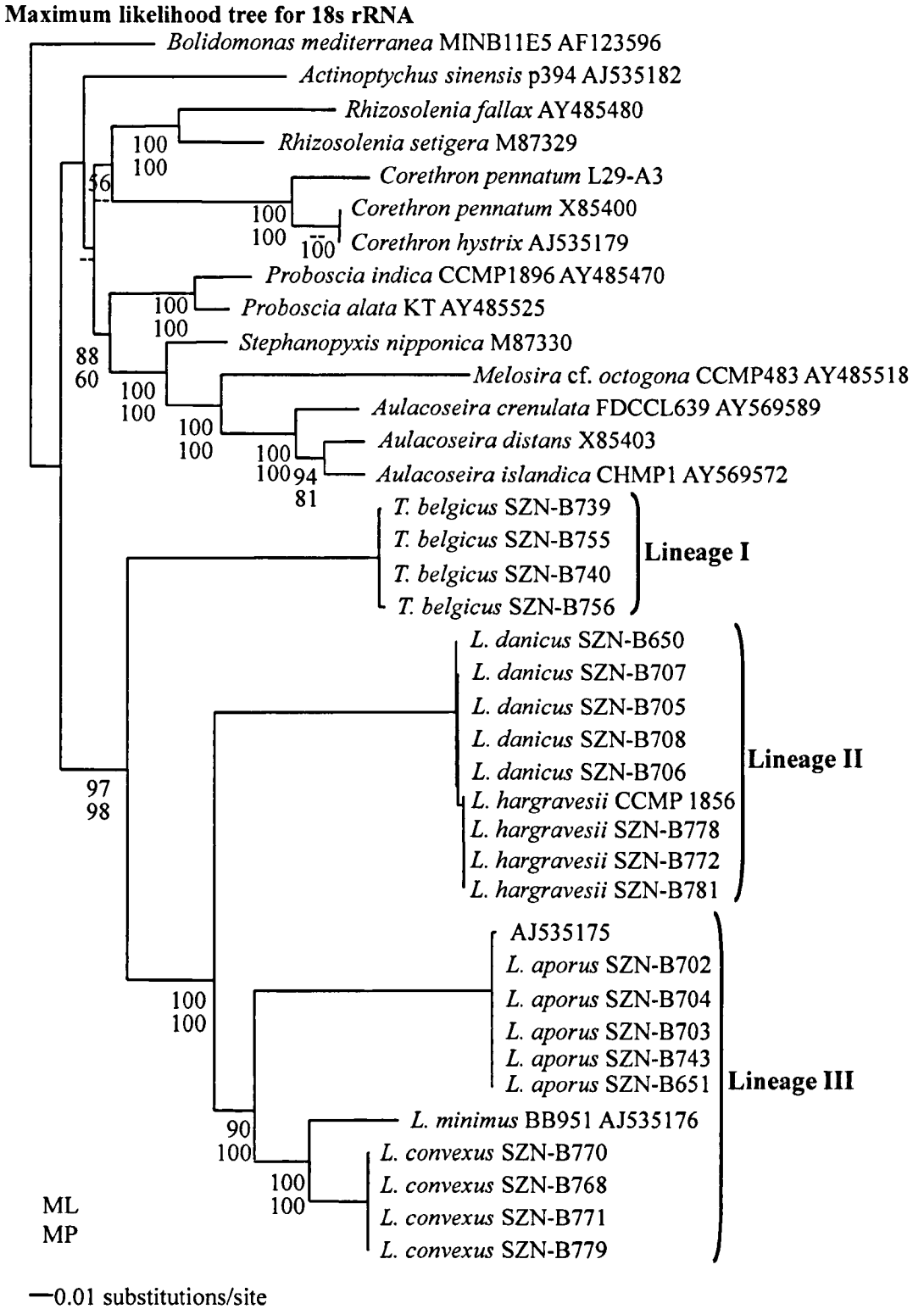


Fig. 2.12. Maximum likelihood tree inferred from 1618 positions of the nuclear SSU rRNA of *Tenuicylindrus*, *Leptocylindrus*, and close groups including *Bolidomonas mediterranea* as outgroup.

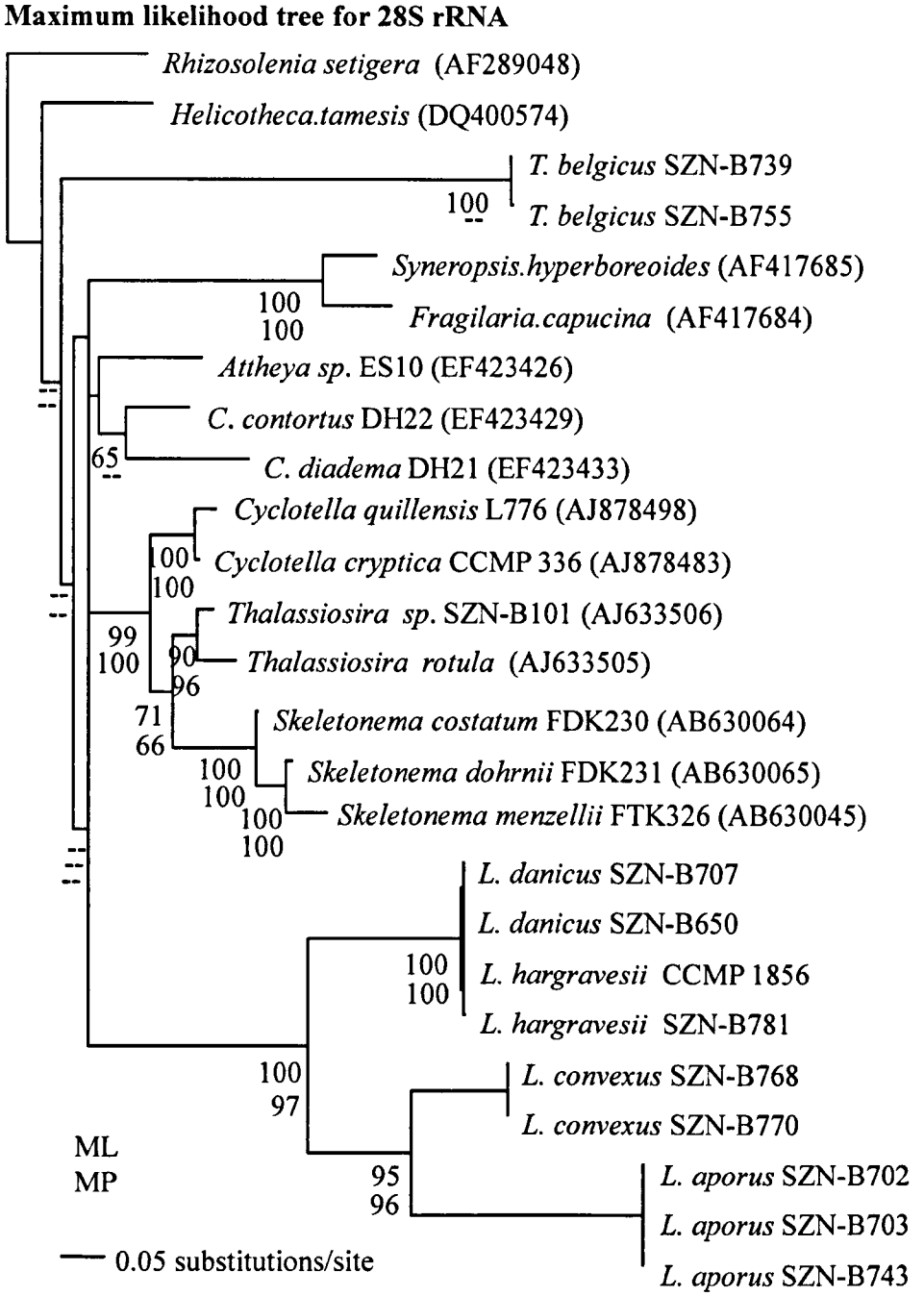


Fig. 2.13. Maximum likelihood tree inferred from 679 positions of the nuclear LSU rRNA of *Tenuicylindrus*, *Leptocylindrus*, and close groups including *Rhizosolenia setigera* as outgroup.

In the plastid SSU rDNA tree (Fig. 2.14), *Leptocylindrus* and *Tenuicylindrus* were resolved close to *Bolidomonas* outgroup. *Leptocylindrus aporus* and *L. convexus* were resolved as sister species, though with low bootstrap support, and *L. danicus* and *L. hargravesii* were recovered together. Bootstrap support for most of the clades in this tree was low or insufficient, meaning that relationships remain unresolved.

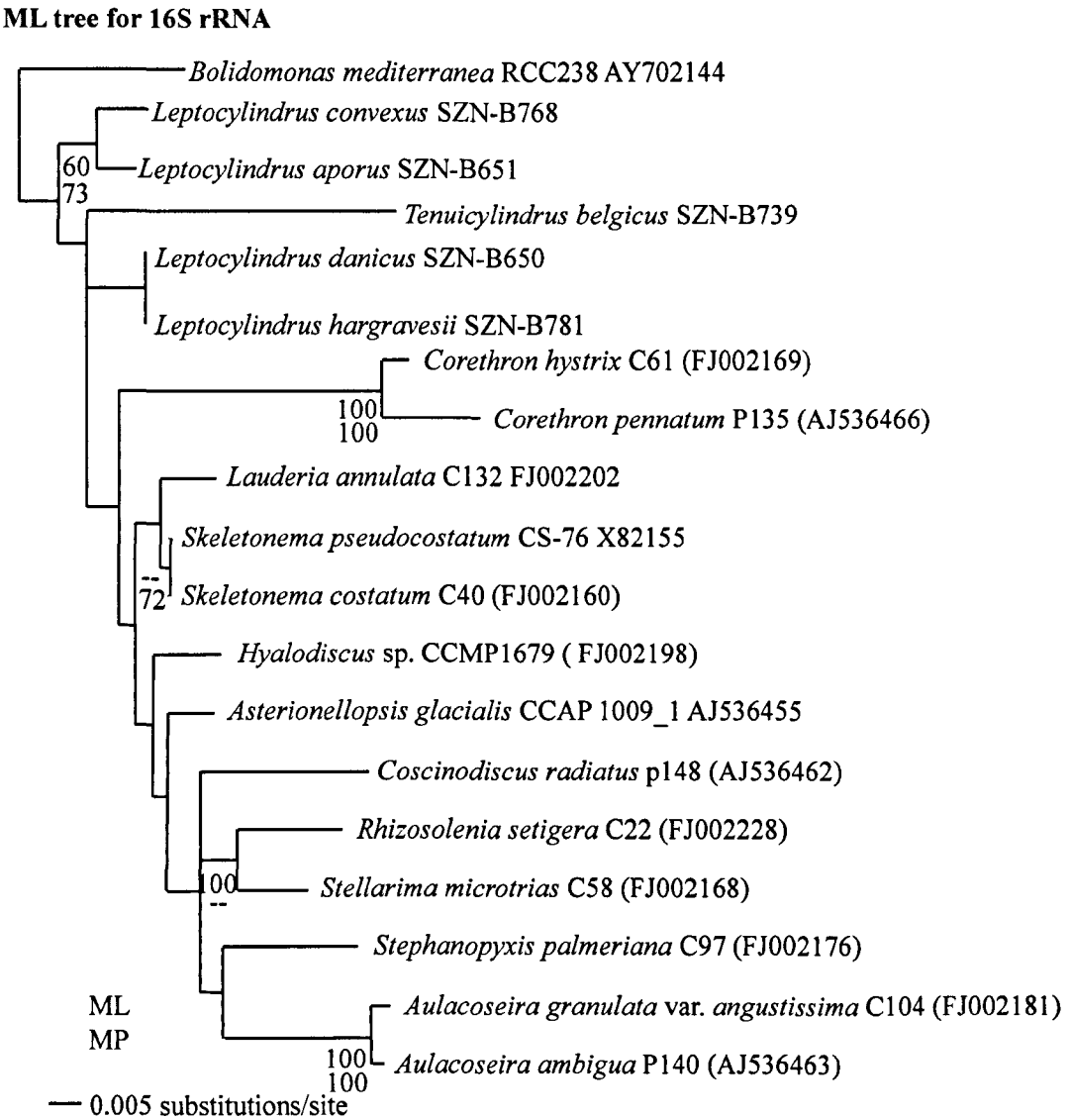


Fig. 2.14. Maximum likelihood tree inferred from 752 positions of the plastid SSU rRNA of *Tenuicylindrus*, *Leptocylindrus*, and close groups including *Bolidomonas mediterranea* as outgroup.

In the psbC tree (Fig. 2.15), *Leptocylindrus* species grouped into a clade. *Tenuicylindrus* was not resolved as sister clade to *Leptocylindrus*, but bootstrap support for the basal clades in this tree was insufficient, thus not ruling out possible sister relationships between the two clades. All clades within *Leptocylindrus* obtained high (100%) to fair (70%) bootstrap support.

In the rbcL tree (Fig. 2.16), *Leptocylindrus* species grouped into a clade with *Tenuicylindrus* resolved as its nearest sister with weak bootstrap support (61–57%). Bootstrap support for the clades within *Leptocylindrus* was insufficient to high.

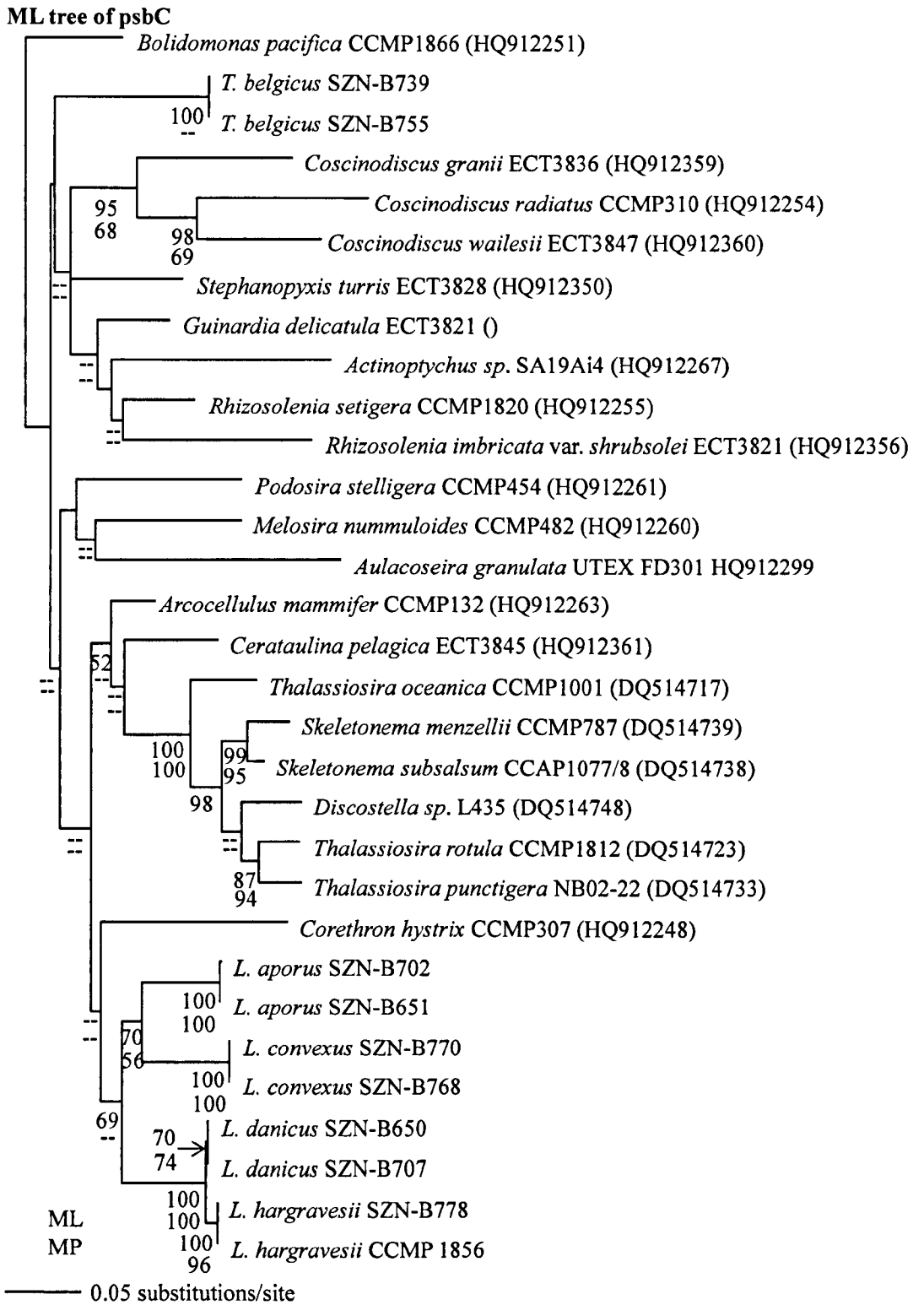


Fig. 2.15. Maximum likelihood tree inferred from 742 positions of the plastid psbC gene of *Tenuicylindrus*, *Leptocylindrus*, and close groups including *Bolidomonas pacifica* as outgroup.

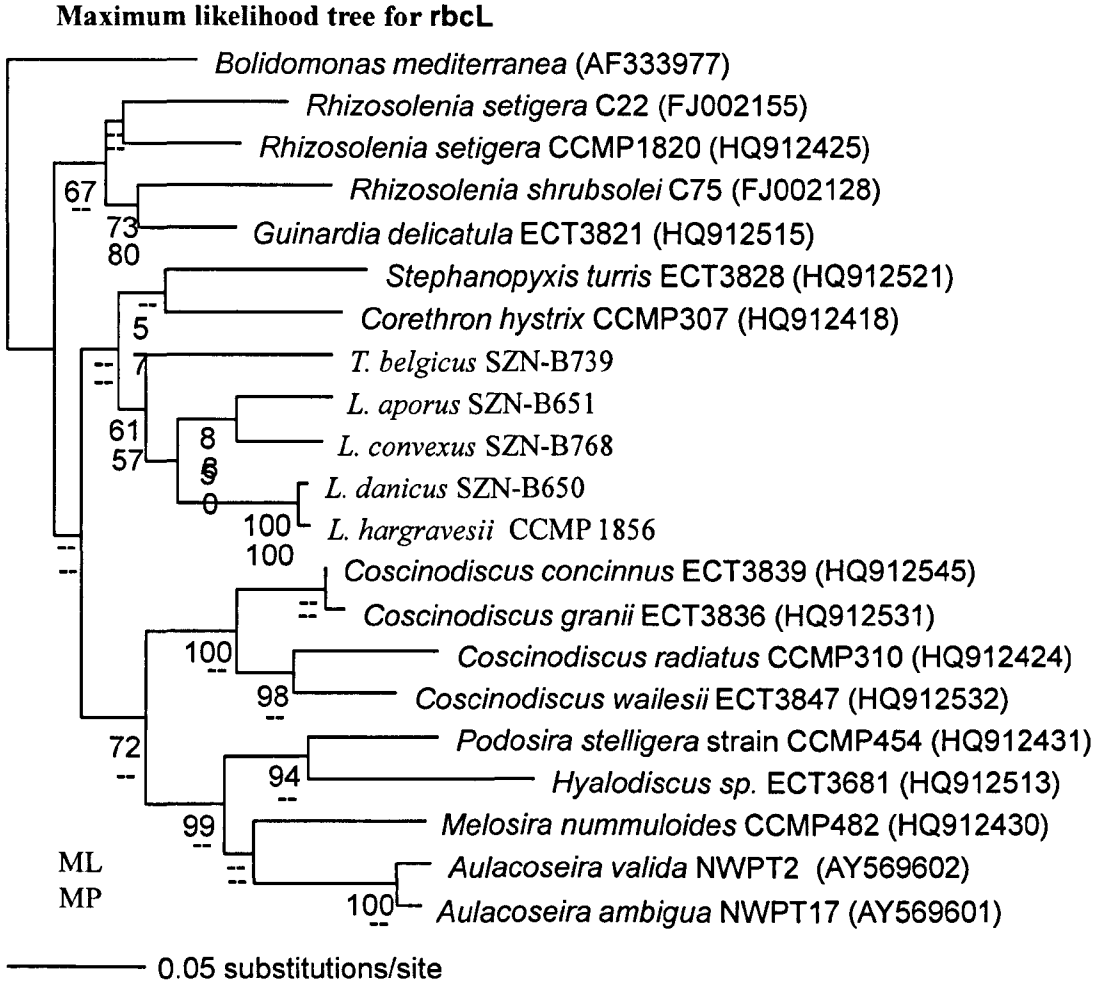


Fig. 2.16. Maximum likelihood tree inferred from 1435 positions of the plastid rbcL gene of *Tenuicylindrus*, *Leptocylindrus*, and close groups including *Bolidomonas mediterranea* as outgroup.

The comparison of the secondary structure of the ITS of the two cryptic species, *L. danicus* and *L. hargravesii*, showed three and one CBC and eleven and thirteen non-CBC in ITS 1 and ITS 2 respectively. The comparison of the secondary structures of different species signifies no conservation in the structure. This means that the species divergence is very ancient and evolution has resulted in the completely different ITS structures for these species.

2.4 Discussion

2.4.1. *Molecular phylogenies* The study on morphological, molecular and life cycle characteristics of species in the genus *Leptocylindrus*, although spatially limited, has

shown that the species *L. danicus*, so far identified in the GoN as one of the most abundant planktonic diatoms, consist in fact of 4 distinct species which can be found over the years in the study area. One of these, *L. aporus*, was considered a variety of *L. danicus*, while two others, *L. hargravesii* and *L. convexus* are new to science. In addition, the present results also indicate that the actual *L. minimus* is not found in GoN, while the species so far identified under this name in the area is related to *L. belgicus*. The latter was so far considered as synonym of *L. minimus* but, in fact shows profound morphological and molecular differences from both *L. minimus* and all the other known species, indicating the need for the establishment of the new genus *Tenuicylindrus* to accommodate it.

The taxonomic history of both genera, *Leptocylindrus* and *Tenuicylindrus*, i.e. of their type species, *L. danicus* and *T. belgicus*, is complicated, which affects to some extent the present definition of the two taxa, prompting some arbitrary and opportunistic choices in the amendment of their descriptions. The taxonomy of *L. danicus* starts with the misrepresentation of the frustule structure by Cleve (1889) based on material collected in his expedition in the Kattegat area (Chapter I). Cleve drew the valves as very convex, only contacting the next ones with a small surface or a point, which is rather unusual for *Leptocylindrus*. In addition, and in contrast with the description, the drawing showed some lines in the cingulum which would indicate that the band margins were at times visible. Later, Cleve (Cleve 1894) himself provided another description of the species as it was, in his words, "originally described from burnt and somewhat misshaped specimens". The preservation error was apparently mended and the re-description was: "*cells cylindrical, with flat (dried convex) ends, forming filaments. Valves without processes or perceptible structure. – Connecting zone very thin, without annuli – Cell contents: a few scattered granular chromatophores. Diameter of the filament 0.01 mm. Length of the cell 0.03 to 0.06 mm*". Gran (1912) undertook a study in the Kattegat bay and described *Leptocylindrus* resting spores as spiny and semi-

circular, and produced through auxosporulation. This peculiarity was confirmed by more recent studies (Hargraves and French 1986) which also described the presence of a subcentral pore in the valve. Therefore, despite the initial uncertainty arising from a type material not representing the species correctly, there are several elements that concur to define the species *L. danicus* in modern terms. However, the finding of at least two very distinct genotypes sharing most of the features used for the definition of *L. danicus* poses further problems as to i) whether separate them as distinct species and ii) which would deserve the name *L. danicus.* One possibility is to consider one single species *L. danicus*, as a genetically diverse taxon. However, the differences among the two genotypes are small for SSU and Chl. 16S but considerable for all the other markers (LSU, ITS, rbcL), which are commonly used in species separation in most diatom genera, and these differences are also accompanied by some subtle morphological differences. For practical purposes, not to keep track of one genotype versus the other might cause confusion in interpreting results of future investigations on the species, e.g. in physiological or 'omic' studies. It appears then more convenient to attribute two distinct names to the two genotypes. The decision about which of the two deserved the name and the designation as neotype of *L. danicus* was arbitrary, though. Unsuccessful attempts were made to retrieve the type material, or material collected in the same cruise by Cleve, in the Swedish Museum of Natural History and Diatom Herbarium, The Academy of Natural Sciences, Philadelphia, USA as well as at the National Botanical Garden of Belgium, where Cleve's materials are maintained. The final choice was hence made based on the higher abundance of one of the two genotypes in the study area, which also contributed to select it for further metabolomics (this study) and transcriptomic analyses (in progress).

It may also be considered arbitrary to establish the genus *Tenuicylindrus* with *T. belgicus* (basionym *L. belgicus*) as the type species, based on the limited information presented in the original description, and with a poor illustration. On the other hand, the

resemblance with the material examined from the GoN is evident in both the drawing and description, especially since the presence of the subvalvar granuli which are typical of the material from the GoN are also perceivable in the illustration by Meunier (1915). By contrast, the designation of the new combination of *L. aporus* based on the description of *L. danicus* var. *apora* is supported by morphological and life cycle characteristics as well as by the presence in GenBank of one sequence obtained from a strain apparently used by Hargraves and French (1986) in the description of this organism (L. Medlin pers. com.), which is identical to the corresponding one of the material examined in the present study.

2.4.2. *Morphological and lifecycle features of* Tenuicylindrus *and* Leptocylindrus *species.* Morphological features in diatom species are historically studied using light microscope and this provides little information of the variation that exists in a species. For this reason, I was interested in studying the genetic diversity of the species. *Leptocylindrus* possesses, perhaps, the simplest known diatom ultrastructure, and until now was considered to comprise of *L. danicus* and *L. minimus*. Vegetative cells of these species are morphologically similar and the main distinction among them is the cell size, plastid number and shape, and spore morphology. The results of the morphological and phylogenetic analyses in this study suggest instead the existence of six taxa associated with the genus *Leptocylindrus*; these can be arranged into three main morphological groups, which correspond to the Lineages I to III identified in the phylogenetic tree. These groups are morphologically and physiologically quite distinct from each other, whereas interspecific differences within the groups are generally more subtle. A comparison between these species is presented in the Table 2.8 and in Fig. 2.17.

The species *T. belgicus* comb. nov., (Lineage I) is morphologically markedly different from *Leptocylindrus* although *L. belgicus* was considered as synonym of *L. minimus*. In

light microscopy the two species are indeed similar in size. *Tenuicylindrus belgicus* possesses two elongated plastids and so does *L. minimus* as described in the literature (Gran 1915, Hasle 1975, Hargraves 1990, Rivera *et al.* 2002). Thus, neither size nor plastid number can be used as distinguishing features in LM. However, the presence of two granules on both side of the cell in *T. belgicus* provides an alternative means to discriminate the species in the light microscope. The ultrastructure of *Tenuicylindrus belgicus* with respect to valve and girdle bands is instead peculiar and quite dissimilar to any of the species in *Leptocylindrus*. The valve mantle of *T. belgicus* is perpendicular to the valve face, while in all *Leptocylindrus* species it lies proximally in the same plane with the valve face, marginally bending and becoming perpendicular to it. The mantle of *T. belgicus* has one free end with a zig-saw format, connecting the other valve mantle, and the opposite end is attached to valvacopula, whilst in *Leptocylindrus* both mantle margins are smooth and attached to the valve face and valvacopula. The short interlinking, triangular processes involved in the tight association with neighbouring cells, have been documented in some fossil records (Crawford and Sims 2008). The intercalary bands in *T. belgicus* have often a lip-like outline with pointed ends, while in *Leptocylindrus* they are nearly trapezoidal or collar-like. All these differences in a range of characters that are rather homogenous across *Leptocylindrus* species support the establishment of a new genus, *Tenuicylindrus*, which is also confirmed by large molecular phylogenetic distances and life cycle data. Indeed another peculiarity of *T. belgicus* is that it maintains a constant cell size throughout the life cycle and hence does not need cell expansion. This kind of constant cell size has also been observed in some other diatoms including for instance *Phaeodactylum tricornutum* (De Martino *et al.* 2007).

The next two lineages are both assigned to the monophyletic genus *Leptocylindrus*. Within this, the two members of Lineage II, *L. danicus* and *L. hargravesii,* are closely related and characterized by the presence of a conspicuous sub-central pore, although

they present subtle morphological differences. The most conspicuous difference concerns the position of the sub-central pore, which in *L. danicus* is adjacent to the

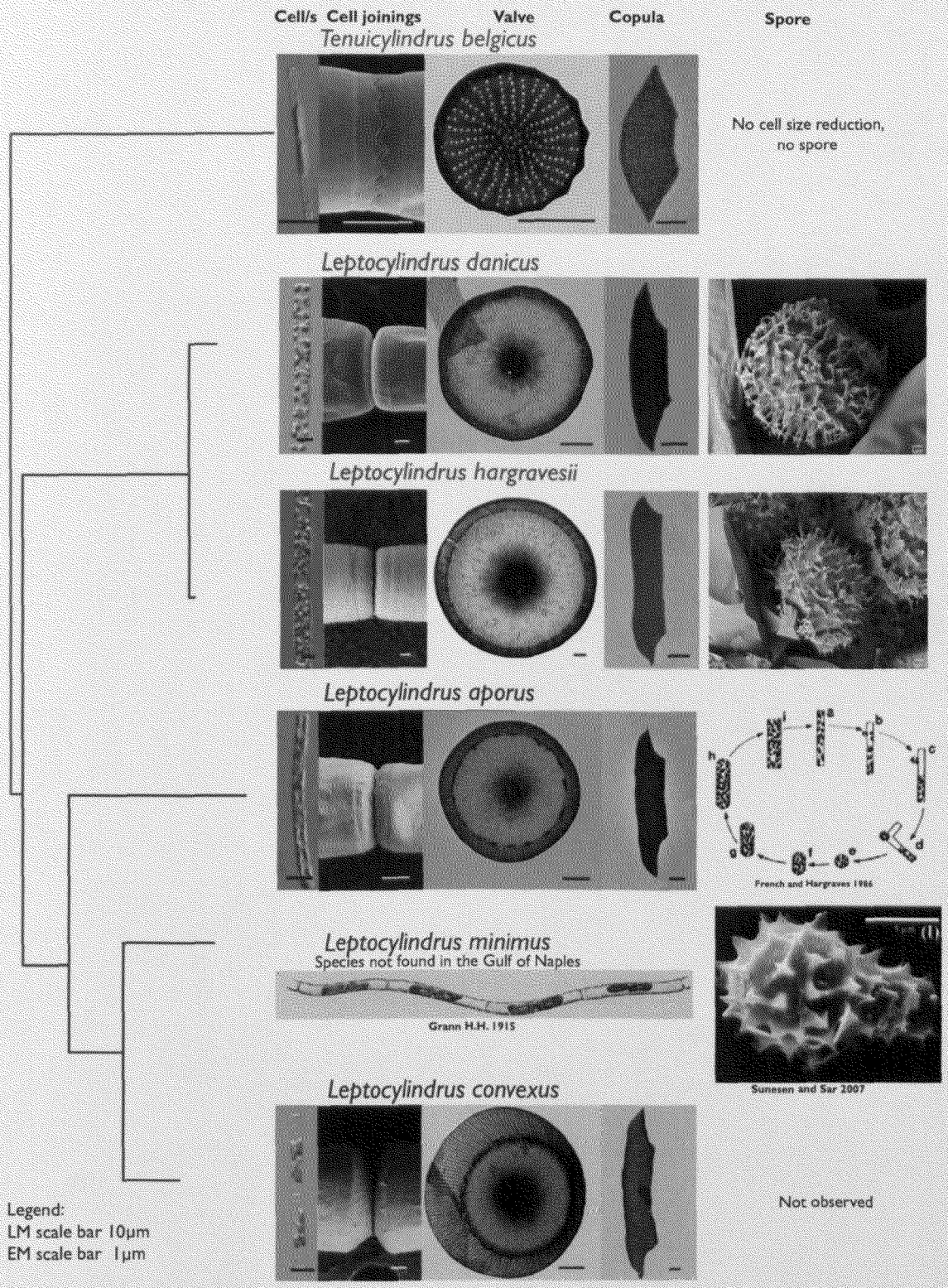


Fig. 2.17. Differential characters for *Tenuicylindrus belgicus* and *Leptocylindrus* species plotted on the SSU rDNA maximum likelihood phylogenetic tree.

Table 2.8. Main morphological characters distinguishing *Leptocylindrus* and *Tenuicylindrus* species.

	L. aporus	*L. convexus*	*L. danicus*	*L. hargravesii*	*T. belgicus*	*L. minimus* Hargraves 1990	*L. minimus* Rivera P. et al. 2002
Cell diameter (µm)	4 – 7.5	3 – 8	3 – 13	3 – 15	2	1.5 – 4.5	2 – 5.2
Cell length (µm)	12.5 – 33	22 – 65	22 – 75	30 – 90	23.6 – 50	-	-
Chloroplast no.	3 – 13	3 – 11	7 – 36	9 – 55	2	1–2	1 – 2
Chloroplast shape	discoid, ovoid	ovoid to elongated	discoid	discoid	elongated	elongated	elongated
Cells per colony	2 – 24	2–68	2 – 165	2 – 162	2 – 14	-	-
Valve to mantle ratio	2.9 – 8.4	2.4 – 4.04	5.6 – 11.45	5.3 – 14.3	-	-	-
Constriction at the cell junction	small	marked	small	small	absent	small	
Auxospores and resting spores	not observed	not observed	semi globular, spiny	semi globular, spiny	not observed	spiny, globular with a cylindrical neck	not described
Subcentral pore on valve	absent	absent	adjacent to circular ring	slightly away from circular ring	absent	absent	absent

hyaline ring in 80% of cells in healthy cultures, whilst in *L. hargravesii*, the pore was never observed adjacent to the hyaline ring. This difference however is difficult to appreciate in spoiled cultures of both species, where the pore position on the valve varies between close to the annulus and valve face border, i.e. adjacent to the mantle. Other ultrastructural differences between the two species concern morphometric characteristics, i.e. the striae and areolae density in the centre of the valve face and the striae density on the mantle are higher in *L. danicus* than in *L. hargravesii*, with scarce or no overlap in the ranges between the two species. The annulus is also larger in *L. hargravesii*, where the centre of the valve appears more densely silicified. Indeed *L. hargravesii* strains appears to be considerably more silicified and stouter than *L. danicus*, and accordingly the samples of the former prepared for electron microscopy retain their morphology better than the latter. In LM, the above mentioned differences are hardly appreciable, but *L. hargravesii* generally has a larger pervalvar axis than *L. danicus*, and it has also a higher number of plastids. However, the number of plastids per unit of cell volume is similar in the two species.

In both *L. danicus* and *L. hargravesii* sexual reproduction was readily induced in monoclonal cultures, producing similar type of auxospores and spores. In all other diatom genera the auxospore produces a vegetative cell of the maximum size (see Montresor, 2006 for a review). The spores produced in *L. danicus* and *L. hargravesii* are similar, but the hypovalve of *L. danicus* has a narrower rim than that of *L. hargravesii*. Even though spores of different size are observed, probably depending on the parent cell size, the cells resulting from spore germination generally show the maximal cell size.

Lineage III includes three species without the sub-central pore, *L. minimus*, *L. aporus*, and *L. convexus*. *Leptocylindrus aporus*, the outermost taxon within the lineage, can be differentiated from other species in the light microscope only based on the relatively smaller size and the shape of the plastids, which are ovoidal as compared to lenticular in *L. danicus* and *L. hargravesii* and more elongated in *L. convexus*. The latter species is

generally larger than *L. aporus* and contrasts with all the other species by its convex valve, which results in a marked constriction at the cell junction level and for the lack of a marked constriction at the cell junction level. Like the other member of this lineage, *L. aporus* does not possess a sub-central pore. Morphological analysis of *L. minimus* was limited because no isolate matched the genetic and morphological features of the species. Based on the available literature, cell diameter and the presence of two plastids match the characteristics of *T. belgicus*, but *L. minimus* has a wider size range of (1.5–5.2 µm; (Hargraves 1990, Rivera *et al.* 2002) as compared to *T. belgicus* (2–2.5 µm, Meunier 1915 and this study). Both species have elongated plastids. The valve structure of *L. minimus* comes closer to the current description of *L. aporus.*

Sexual reproduction and spore formation were not observed in the species of lineage III. However vegetative cell size expansion through auxospore-like structures was observed in *L. aporus*, in agreement with the observations of French and Hargraves (1986). A vegetative mechanism of cell size expansion was initially reported by von Stosch (1965), followed by similar observations on a number of centric diatoms, for example, *Skeletonema costatum*, (Gallagher 1983), *Coscinodiscus wailesii* (Nagai *et al.* 1995), as well as for the pennate diatom, *Achnanthes longipes* (Chepurnov and Mann 1999). The vegetative mechanism is believed to have a selective advantage because of the lower energy requirement and because it overcomes the risk of finding a mate of the opposite mating type. As an autapomorphic character, *L. minimus* only is reported to produce resting spores with a different morphology, but whether they also develop from a sexual auxospore is unclear. This latter characteristic, along with the number and shape of the plastids, could also help discriminating this species from *L. minimus*, which instead do form spiny globular spores with a neck shaped structure at the basis (Hargraves 1990). Spores were not produced in *L. convexus* cultures either, but a rather strange silicified structures of maximal cell size were observed (data not presented), which requires further investigations.

2.4.3. ***Variation of plastid morphology and number.*** The species of Leptocylindraceae show differences in plastid size, number and morphology. The number of plastids varies from two per cell, in *T. belgicus,* to many per cell as in *L. danicus* and *L. hargravesii.* The plastids can be ovoid as in *L. danicus* and *L. hargravesii* to very elongate as in *T. belgicus*, with some intermediate forms as in *L. aporus* and *L. convexus.* These differences can result from environmental factor such as light regime, as well as intrinsic factors such as cell size relative to plastid size. Light induced changes in plastid shape if a species will especially affect light-harvesting and -utilisation, and can be observed as changes in the pigment content and intracellular self-shading. Characteristically, low light acclimated cells have plastids evenly distributed in the cells (maximising light harvesting), while high light acclimated cells have condensed plastids (less light-absorbing surface) (Blatt *et al.* 1981).

2.4.4. ***Temporal and spatial distribution of the species.*** With its simple morphology, the species has been considered widespread, mostly coastal, with numerous records of its occurrence and often being the major contributor of the diatom bloom in many parts of the ocean. *Tenuicylindrus belgicus* (formerly *Leptocylindrus belgicus*) has never been reported elsewhere after its first report by Meunier (1915) from the Southern North Sea, Belgium. However, this may be due to a lack of recognition of this taxon as a species different from *L. minimus.* Indeed, the species illustrated in Round *et al.* (1990, 342, Fig. a) as *L. minimus* clearly shows the characteristic features that are identified in *T. belgicus.* Unfortunately, it is not possible to determine the geographic origin of the specimen illustrated in that picture. Other specimens attributable to *T. belgicus* are shown in the manual by Kraberg *et al.* (2010, page 82, figures a and b). Meunier (1915) and Kraberg *et al.* (2010) reported an early spring-summer period of occurrence for the species, whereas in the GoN the species, easily identified in LM and so far classified as

L. minimus, is observed from late summer to autumn period at our LTER-MC station, with a narrow period of bloom in autumn (Ribera d'Alcalà *et al.* 2005).

For the *Leptocylindrus* spp., the seasonal distribution is mainly based on the number of clones recovered from the cell brought under laboratory cultivation, as the species have not been discriminated in morphological observations at the LTER-MC stations so far. The distribution of *L. danicus* is wider than the other species recognized in the study, with its population recorded from late autumn through mid-summer, while *L. hargravesii* has the narrowest distribution (December–January) over the year in the GoN. *L. hargravesii* is very rarely brought into cultures; hence they may be rare in natural populations or, alternatively, is difficult to be grown in laboratory conditions, as it happens with many phytoplankton species. *Leptocylindrus aporus* is apparently the species responsible for the remarkable summer blooms so far attributed to *L. danicus*, and it is also widely found in autumn, along with *T. belgicus* seems to be mainly a summer blooming species and *L. danicus* is present in all the seasons except summer, other species are also not found during summer. Thus *L. aporus* seems to have evolved to bloom during warm season. *Leptocylindrus convexus* also occurs mostly in the winter season from December to March but was never found in very high numbers. *Leptocylindrus minimus* was not found in the GoN, but the possibility that it escaped sampling cannot be totally excluded.

A more rigorous and detailed study on the autecology of the species needs to be made through the application of sequence based techniques including Fluorescent In Situ Hybridization (FISH), quantitative-PCR (qPCR) or construction of clone libraries with species specific primers. With respect to the spatial distribution of the recognized species, based on the metagenomic sequences available in GenBank has revealed that the species are not restricted to the study site but occur from the South China Sea to the north Atlantic Ocean (Chapter III).

2.5. Conclusion:

The study provides a description of the genetic diversity of the genus *Leptocylindrus* in the GoN. In addition to *L. danicus*, which was considered to be the most abundant species in the GoN, two new species, *L. hargravesii* and *L. convexus*, were found. The variety *L. danicus* var. *aporus* probably is a distinct species. Finally, the species so far identified as *L. minimus* in the GoN proved to be very different from any *Leptocylindrus* from the ultrastructural and molecular point of view and was recognised as synonym of the species *L. belgicus*. Due to the marked differences observed for this species, the establishment of the new genus *Tenuicylindrus* was proposed to accommodate it.

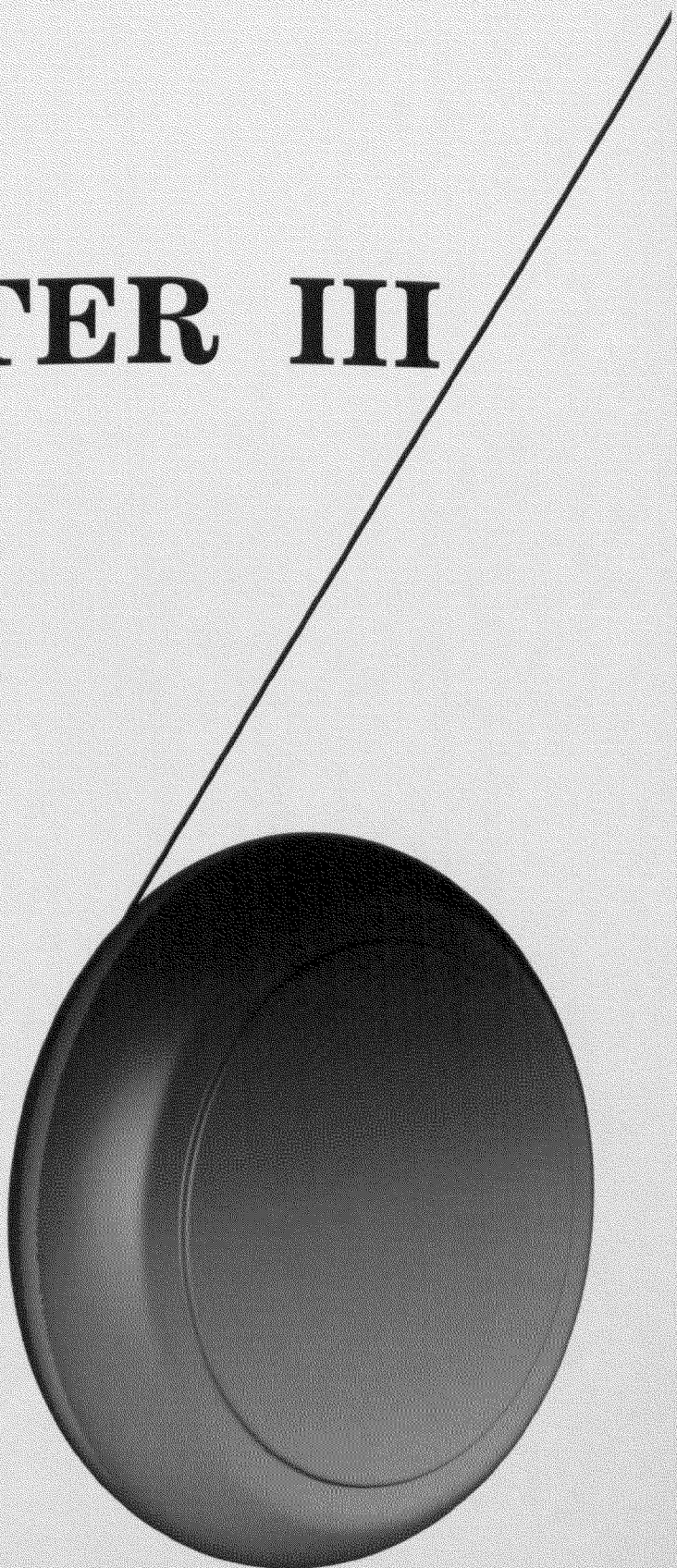
It is relatively easy to distinguish the five species in light microscopy, but in natural samples this is not always possible. The often subtle morphological differences that exist for light microscope differentiation among the currently recognized *Leptocylindrus* species were detected in laboratory cultures, whereas more information would be needed on the natural variability of the different characters. It is often impossible to discriminate, for example *L. danicus, L. hargravesii* and *L. aporus,* in case plastid shape is not well preserved. In such cases electron microscopic examination may permit identification, and in some cases this may need confirmation from molecular analyses. New molecular methods like Fluorescent In Situ Hybridization (FISH) would be simple and convenient to identify in natural samples.

This study shows that also in *Leptocylindrus* it is possible to detect species new to science. In fact, species diversity in this genus (and its sister Genus *Tenuicylindrus*) has increased from two to five. This is a state of affairs similar to that described for the increase in species diversity in the polar centric genus *Skeletonema* according to Sarno et al. (2005, 2007). Species diversity may increase further if samples from other regions are screened for *Leptocylindrus* and *Tenuicylindrus* species, comparable to what was found in the global biogeographic study of *Skeletonema* by Kooistra et al. (2008). Nevertheless, if *Leptocylindrus* and *Tenuicylindrus* constitute the sister clade of a clade

with all other diatoms, then even the discovery of a series of additional species in these two genera will leave the clade as a species poor group in comparison to the estimated 200,000 species in the rest of the diatoms.

Tenuicylindrus and *Leptocylindrus* species appear to have evolved very distinct life cycle mechanisms, encompassing no cell size reduction along the vegetative cycle, vegetative autoenlargement, spore formation with no apparent sexual mechanisms, and spore formation following sexual reproduction. Whether the differences in the life cycle mechanisms have influence on the success of the species and the extent to which these complex mechanisms have helped the species in the adaptation to their environment needs to be studied. Results of such studies would help in gaining insight into the ecology of the species.

CHAPTER III



Molecular detection and distribution of *Leptocylindrus* species using metagenomic databases

3.1 Introduction

The assessment of biogeography of diatom species started with the comparison of species and the mapping of their distribution patterns by Ehrenberg in 1843 in his "Microgeologie." Further studies included those by De-Toni, "Sylloge Algarum" and by Cleve, who introduced the term 'plankton type' to classify and categorise plankton species and to track their distribution. Smayda (1958) and Braarud (1962) emphasised the ecological and biological factors that govern the distribution of phytoplankton species. Initially, biogeographic studies were based almost exclusively on species as defined by frustule morphology. This approach has often resulted in overly inclusive taxonomic units. On the other hand, over-interpretation of environment-induced, subtle morphological differences among allopatric populations has led to erroneous descriptions of new taxa. For example, *Thalassiosira gravida* Cleve and *T. rotula* Meunier were originally considered to be two distinct species, the former being restricted to high latitudes and the latter to lower latitudes (Hasle 1976). However, culture experiments with decreasing temperature, showed that valve morphology changes from typical *T. rotula* to typical *T. gravida* (Syvertsen 1977). Despite these artefacts and shortcomings, diatom classification continues to rely on the morphological distinction as the method of choice for species delimitation and has resulted in the description of over 25,000 species (Alverson 2008).

Recent studies have shown that diatom species diversity has been severely underestimated. Discontinuous morphological variation coupled with variations in reproductive, molecular, genetic, physiological, and ecological characters exist within morphologically delineated species (species sensu lato). For example, *Skeletonema costatum* sensu lato was found to include several morphologically, ecologically and genetically different taxa (Sarno *et al.* 2005, Sarno *et al.* 2007). *Pseudo-nitzschia delicatissima* and *P. pseudodelicatissima* were

found to consist each of several reproductively isolated taxa, with subtle morphological differences amongst them (Amato 2007, Amato 2008, Lundholm 2012). Similar examples of such diversity within *Pseudo-nitzschia* species *sensu lato* were reported by (Lundholm 2002, Behnke 2004, Mann 1999, Montresor 2003, Šlapeta 2006).

The diversity in a particular geographical region is the result of interactions between processes that add species, such as allopatric species formation and geographic dispersal, and processes that lead to local extinction, such as predation, competitive exclusion, and stochastic variation. While all these processes regulate distribution of macroscopic organisms, their relevance to microorganisms is still debated. Conceptually, microorganisms, including prokaryotes, unicellular eukaryotes, and small multicellular organisms, are assumed to be cosmopolitan, the hypothesis, "everything is everywhere, but the environment selects" (Baas-Becking 1934, Fenchel 2006). This view has been empathised in studies applying ecological methods and concepts on both original studies and literature data of heterotrophic protists, mainly ciliates (Fenchel 1997, Finlay 2002, Azovsky 2002). The most commonly suggested elucidation for cosmopolitism is that large population sizes, short generation times (Fenchel 2004, Coleman 2002, Finlay 1999) and the ability to form dormant stages (cysts, eggs and spores) facilitate dispersal. Water movement interconnects water masses at various geographic scales (Cermeño 2009) and therefore, allopatric speciation should be rare or non-existent (Finlay 2002). In this hypothesis, the distribution of microorganisms is restricted only by adverse environmental conditions (Baas-Becking 1934)

The alternate hypothesis states that dispersal capability is low, and that restricted geographical distribution patterns promote speciation. Recently, studies have demonstrated that microorganisms exhibit biogeographic patterns, but that the rates of the underlying processes vary more widely for microorganisms than for macroorganisms (Whitaker *et al.* 2003, Martiny *et al.* 2006, Telford *et al.* 2006). One approach for testing cosmopolitism is

through the study of similarities or dissimilarities in species composition among geographically distant sites with similar or different environmental characteristics (Green and Bohannan 2006).

The advancement of large scale sequencing capacity has enabled a comprehensive view of the microbial diversity. New technologies, like metagenomics, which can quantify the total genetic diversity including unculturable species, have opened new possibilities to investigate the hidden diversity, triggering geographic distribution studies of phytoplankton. The sequence data can be utilized to quantify the degree of overlap in microbial assemblages between habitats and to estimate (or roughly approximate) the richness of microbes in the world's oceans and coastal seas. In addition massive sequencing data provide a powerful tool to describe molecular species diversity, and may discover rare species in samples, discriminating between the two alternative hypotheses of species distribution.

Recent metagenomic projects apply next generation sequencing (NGS) technologies such as Illumina, SOLiD, PacBio and Ion Torrent, which are high throughput and economic than the traditional Sanger sequencing technology. These have led to numerous metagenomic projects in diverse environments such as the human microbiome (Group *et al.* 2009), and virome (Dinsdale *et al.* 2008). Sequence based surveys based mainly, in the most common being small subunit rRNA-based surveys, rely on the resolution provided by small subunit rRNA (SSU; in eukaryotes 18S rRNA and in prokaryotes 16S RNA) have led to the discovery of important groups such as the third branch on the tree of life (the Archaea) (Woese and Fox 1977). Many of these newly discovered lineages elude discovery in the laboratory because they cannot be cultured in the laboratory. The 18S rRNA of eukaryotic organisms is typically 1750 bp long and can be divided into nine variable parts, V1–V9, according to Nelles et al. (1984), that provide an appropriate phylogenetic signal for genealogy and systematics (Medlin and Kooistra 2010). The

presence of highly conserved regions in the SSU rDNA gene at several positions has facilitated designing and use broadly targeted oligonucleotide primers that work on a wide diversity of species for both sequencing and amplification by the polymerase chain reaction (PCR). Two hypervariable regions of SSU, the V4 and V9, have been used recently as targets for high throughput sequencing (Amaral-Zettler 2009, Stoeck 2009) and discrimination of thousands of phylotypes (or operational taxonomic units, OTUs) that serve as a provisional surrogate for "species", in metagenomic studies. These groupings have revealed that the level of DNA diversity is much higher than previously known, especially in marine planktonic microorganisms (López-García and Moreira 2008). However, even V4 or V9, is considered to be too conservative to distinguish between closely related species within some taxa (Stoeck *et al.* 2010).

Despite the advantages, metagenomic projects engaging current and emerging NGS technologies, face analytical challenges because of the need to balance the desire for high quality methods with the need for automation to keep up with the sequence onslaught (Wu *et al.* 2008). Firstly, the SSU rDNA sequence occurs in multiple copies, and in theory, these copies can show base differences. This means that a population may exhibit micro-variation, i.e., huge numbers of highly similar sequences without any phylogenetic structure. Second, sequence-based surveys through metagenomic projects, studying diversity in the environmental samples produce massive amounts of information (due to artificial PCR-based duplication of reads, (Niu *et al.* 2010) that are difficult to organise, integrate and finally to classify species. This wealth of information is rarely mined to its full potential. Third, sequence errors and artefacts produced are more than traditional Sanger sequencing, which is further propagated in the analysis leading to false elucidations (Gomez-Alvarez *et al.* 2009). Finally, the sequence diversity may vary over time (season) and the sample reflects the diversity found only during that moment. Therefore, it is

necessary to record this information with as much detail as possible that can be useful to different user needs and subsequently be interpreted with best possible way.

Clustering analysis, i.e. methods that identify and cluster groups of similar objects, can be applied effectively to resolve the above-mentioned challenges (Li *et al.* 2012). By sequence clustering a large redundant dataset can be represented with a small non-redundant (NR) set, which requires less computation. Clustering also helps identifying errors by using a consensus from sequences and thus leads to an efficient classification of species. There are several programs available for clustering but the recent ones are fast programs and hence popular. These include CD-HIT, first fast and comprehensive clustering package (Huang *et al.* 2010), others include DNACLUST (Ghodsi *et al.* 2011), Uclust (Edgar 2010), and SEED (Bao *et al.* 2011). The choice of the program mainly depends on the user's requirements and preferences, for a review on different methods see (Li *et al.* 2012).

In a previous study (Chapter II) we obtained monoclonal strains from single chains isolated from net samples taken in the Gulf of Naples (GoN). Through this study, the morphology and biology of the species has been associated to the type sequence has been made, but this is restricted to few observations being made at a single geographical location, which renders this approach not effective in capturing the total diversity. Instead, massive nuclear rRNA encoding sequence data bases obtained from NGS approaches as of environmental samples allow to obtain a far more extensive sample of the diversity, but it is based exclusively on the obtained sequences at hand.

Using metagenomic data from distant geographical regions and contrasting ocean environments, we surveyed the spatial distribution of marine diatom species of the genus *Leptocylindrus*. Owing to the relatively contrasting or similar physicochemical conditions between sites in the ocean, our analysis, perhaps, also allows addressing the eco-physiological characteristics of the species. The radial centric diatom genus *Leptocylindrus*

is an example of pseudo cryptic diversity where gross intraspecific morphological variation and minor morphological differences among species mask actual genetic diversity (Chapter II). As an example, *Leptocylindrus danicus* is a common species in the coastal environments of temperate to sub-tropical regions, and is recorded also in the Antarctic and Artic regions, where it is often prominent species during diatom blooms. The species has a simple morphology and hence was considered single species. Other species in the genus included, mainly, *L. minimus* comparatively less abundant and less widespread (Hargraves 1990).

A detailed investigation of the species, *L. danicus*, genetics and frustule morphology has led to the discovery of an additional three species. First, *L. hargravesii*, a close relative of *L. danicus,* showed a similar morphology and life cycle pattern. Thereafter two morphologically and physiologically distinct species, *L. aporus* and *L. convexus,* were described (Chapter II). In the GoN, in some cases the species occurrence overlapped in time, and in some others exhibited an opposite pattern of occurrence (Chapter II). Hence, it is intriguing to understand the biogeographic distribution of each of these species across the Ocean.

In the present study, the following questions were addressed:

1. Whether the marker sequences within clades of *Leptocylindrus* species *sensu stricto*
    a) show high micro-variation (many end branches) or not,
    b) consist of one or a few end branches with large numbers of copies and the remaining end branches representing single copies, or if there are many end branches within a clade containing multiple copies,
    c) are on long internal branches (are many base pairs different from their sister sequences) or not,
2. Whether the marker sequences of newly recognized *Leptocylindrus* species in Chapter II are each recovered in their own clade of environmental sequences, or if

marker sequences of multiple species are within single clades without any internal phylogenetic structure.

3. Whether additional diversity exists in *Leptocylindrus* and *Tenuicylindrus*, i.e., there are well delineated clades firmly embedded within *Leptocylindrus* and *Tenuicylindrus* without any sequences of the species *sensu stricto* delineated in Chapter II,
4. Whether the *Leptocylindrus* species *sensu stricto* that pass the test of being real entities, are widespread or occur only locally,
5. Whether the species *sensu stricto* within the *L. danicus* complex are as widespread as the historically recognized *L. danicus sensu lato*,
6. Under what ecological conditions the species are found.

To address these questions, I explored metagenomic databases that were generated from DNA and RNA extracted from geographic water samples gathered in the projects "Biodiversity of Marine Eukaryotes" (BioMarKs) and *Tara* Oceans. BioMarKs is a European Union ERA BioMarKs project involving experts from eight EU research institutes studying eukaryotic microbial taxonomy and evolution, marine biology and ecology, genomics and molecular biology, bioinformatics, as well as marine economy and policy. The "*Tara* Oceans expedition" is three-year global ocean expedition to study the impact of climate change on the microscopic life forms in the ocean. Species diversity was recorded using variable regions of the nuclear SSU rDNA. The databases have been developed with the goal to record the diversity and influence of climate change on the microscopic life forms.

Note that the sampling scheme was one of opportunity and that a species may be present at a site but remained undetected because it was below detectable concentrations at the time of sampling.

3.2 Methods

3.2.1 Sampling and databases. The "BioMarKs" project assessed biodiversity of unicellular eukaryotes at nine different European marine stations through development of a SSU rDNA database (Table 3.1). For the present study, samples from six stations were considered: Oslo Fjord (sampled; Sept 2009 and June 2010), Roscoff (the English Channel; sampled April 2010), Gijon (Bay of Biscay; Sept. 2010), Barcelona (Mediterranean Sea; Feb. 2010) and Naples (Mediterranean Sea; Oct. 2009 and May 2010) and Varna (Black Sea; May 2010). In the BioMarKs project both the V4 and V9-regions were sequenced, but in the present study, only the V4 region was scanned for the presence of *Leptocylindrus*-like sequences.

For protocols on sampling and sequencing refer to BioMarKs reports on website (http://www.biomarks.eu/) and refer to Logares *et al.* (2012). Briefly, seawater samples were taken with Niskin bottles from the subsurface (1m) and to the depth of deep chlorophyll maximum (20m, DCM) or to the sediment. Subsurface and DCM samples were fractioned into 0.8–3µm, 3–20 µm and 20–2000µm through polycarbonate filters (142 mm in diameter). Filters were flash frozen and stored at -80°C. Sediment samples were taken with sediment cores and small aliquots were frozen at -80°C for downstream molecular analysis.

Total DNA and RNA were extracted from the same filter using the NucleoSpin RNA kit (Macherey-Nagel, Hoerdt, France) and from sediment samples were extracted using the RNA Power Soil Total Isolation kit combined with DNA Elution Accessory kit (MoBio Laboratories). Extracted RNA was reverse transcribed to DNA using the RT Superscript III random primers kit (Invitrogen, Carlsbad, CA, USA) and universal primers for the V4 region (Stoeck *et al.* 2010).

The PCR mixture (25 µL final volume) contained 5 ng of template with 0.35 µM of each primer, 3% of DMSO and 2X of GC buffer Phusion Master Mix (Finnzymes).

Amplifications of the V4 region were done in triplicates following the PCR program: initial denaturation step at 98 °C for 30 sec, followed by 10 cycles of 10 sec at 98 °C, 30 sec at 53 °C, 30 sec at 72 °C, followed by 15 cycles of 10 sec at 98 °C, 30 sec at 48 °C, 30 sec at 72 °C and final elongation step at 72°C for 10 minutes. Amplicons were then pooled and purified using the NucleoSpin® Extract II kit (Macherey-Nagel, Hoerdt, France). Amplicon pools where finally sequenced at the CEA Genoscope in Evry using a GS FLX emPCR Genomic Lib-L kit according to the manufacturer's protocol (Genome Sequencer FLX Titanium, 454 Life Sciences from Roche, Brandford, CT, USA). The "*Tara*-Oceans" project assessed biodiversity of unicellular eukaryotes at a large series of marine sample stations, the DNA harvested at 35 of these have been considered in the present study. These stations were located in the Mediterranean Sea, The Red Sea, the Arabian Sea, the wider Indian Ocean, Cape Town, Southern and Central Atlantic Ocean, the Antarctic Peninsula, and the eastern side of the southern and central Pacific Ocean (Fig. 3.1, A and B). This project sequenced only the V9 variable region of SSU rDNA as barcode for documentation of species diversity and abundance across continents and different biomes. The colossal database was explored for sequences of *Leptocylindrus* species.

For protocols on sampling and sequencing refer to Karsenti *et. al.* (2011). Briefly, seawater samples were taken with Niskin bottles at surface and DCM were sequentially obtained in four size fractions, 0.8–5µm, 5–20µm, 20–180µm, and 180–2000µm. Filters were flash frozen and stored at -80 °C. Sediment samples were taken with sediment cores and small aliquots were frozen at -80 °C for downstream molecular analysis. DNA extraction was done following the protocol of BioMarKs. If the sample contained low amount of DNA then the whole genomic DNA was amplified. The PCR mixture was the same as that of BioMarKs protocol. PCR program for amplification of V9 fragment involved: initial denaturation step at 98°C for 30 sec, followed by 25 cycles of 10 sec at 98 °C, 30 sec at 57 °C, 30 sec at 72 °C, followed by 15 cycles of 10 sec at 98 °C, 30 sec at 48 °C, 30 sec at 72

°C and final elongation step at 72 °C for 10 minutes. Triplicate reactions were pooled and the pooled amplicons of the V9 fragment were sequenced using a Genome Analyser IIx system (Illumina, San Diego, CA, USA).

3.2.2 Sequence Retrieval. Sequences were retrieved from already defined Bacillariophyta datasets of the V4 region from the BioMarKs team (courtesy of Dr Stephane Audic) and the V9 region from the *Tara* Oceans team (courtesy of Dr Chris Bowler). The query-datasets were clustered with the V4- and V9-regions in the nuclear SSU rDNA sequences of 134 diatom species available at the LEEP laboratory at SZN (reference dataset), as reference for clustering, with the CD-HIT-EST-2D module of CD-HIT suite (http://weizhong-lab.ucsd.edu/cdhit_suite/cgi-bin/index.cgi). Reference sequences of each of the *Leptocylindrus* species and *Tenuicylindrus belgicus* (Chapter II) were included in the reference dataset.

Pairwise distances among *Leptocylindrus* species, *Tenuicylindrus belgicus* and outgroup taxa were computed for the V4 and V9 regions using PAUP* (Phylogenetic Analyses Using Parsimony and other methods; version 4.0b10) (Swofford 1998) to define the sequence identity (similarity) cut-off value for clustering of query sequences with the reference sequences. A value too close to 100% would result in finding only sequences almost identical to the reference sequences of the known species, whereas setting the value too low would not only include sequences belonging to possible yet-unknown clades within *Leptocylindrus* and *Tenuicylindrus*, but also include possibly thousands of sequences outside these genera.

The sequence identity (similarity) cut-off value was set at 0.90. All sequences in the query dataset with similarity $\geq$0.90 to any of the 134 reference sequences were gathered from the query datasets and matched with the most closely related reference sequence. In addition,

Table 3.1. Metadata for the BioMarKs sampling stations.

Session	Depth	Location of stations		Name	Date	Temperature (°C)		Salinity (psu)		Chl a (μg L^{-1})
		Lat.	Lon.			Sur.	Bot.	Sur.	Bot.	
Oslo 2009	Pl.	59.25	10.71	–	22 Sep, 2009	15.5	8.0	25.0	35.0	3.0
Oslo 2009	Sed. A	59.25	10.71	Station A 103m depth	23 Sep, 2009	15.5	8.0	25.0	35.0	3.0
Oslo 2009	Sed. B	59.26	10.72	Station B 24m depth	23 Sep, 2009	15.5	8.0	25.0	35.0	3.0
Naples 2009	Pl.	40.81	14.25	LTER MareChiara	13 Oct, 2009	22.8	14.6	37.7	37.9	–
Naples 2009	Sed.	40.81	14.25	LTER MareChiara	15 OCT, 2009	22.5	14.6	37.7	37.9	–
Barcelona 2010	Pl./Sed.	41.67	2.80	Blanes Bay	9 Feb, 2010	12.5	12.5	37.5	37.8	1.3
Roscoff 2010	Sed.	48.77	−3.96	SOMLIT Astan	Pl 20 Apr, 2010/Benthic 26 Apr, 2010	10.0	10.0	34.9	34.9	–
Roscoff 2010	Sed.	48.77	−3.96	SOMLIT Astan	Pl 20 Apr, 2010/Benthic 26 Apr, 2010	10.0	10.0	34.9	34.9	–
Naples 2010	Pl./Sed.	40.81	14.25	LTER MareChiara	Pl 11 May, 2010/Benthic 12 May, 2010	19.2	14.5	37.1	37.9	1.2
Oslo 2010	Pl.	59.25	10.71	–	22 JUN, 2010	15.0	6.0	–	–	–
Oslo 2010	Sed.	59.25	10.71	Station A 103m depth	23 JUN, 2010	15.0	6.0	21.5	34.5	–
Varna 2010	Pl./ Sed.	43.17	28.83	–	26–27 JUN, 2010	21.5	–	–	–	–
Gijon 2010	Pl./ Sed.	43.67	−5.58	S–2 of Radial de XiXón	14 SEP, 2010	20.0	–	–	–	–

Lat.– Latitude, Lon.– Longitude, Pl.– Plankton, Sed.– Sediment, Sur.– Surface, Bot.– Bottom

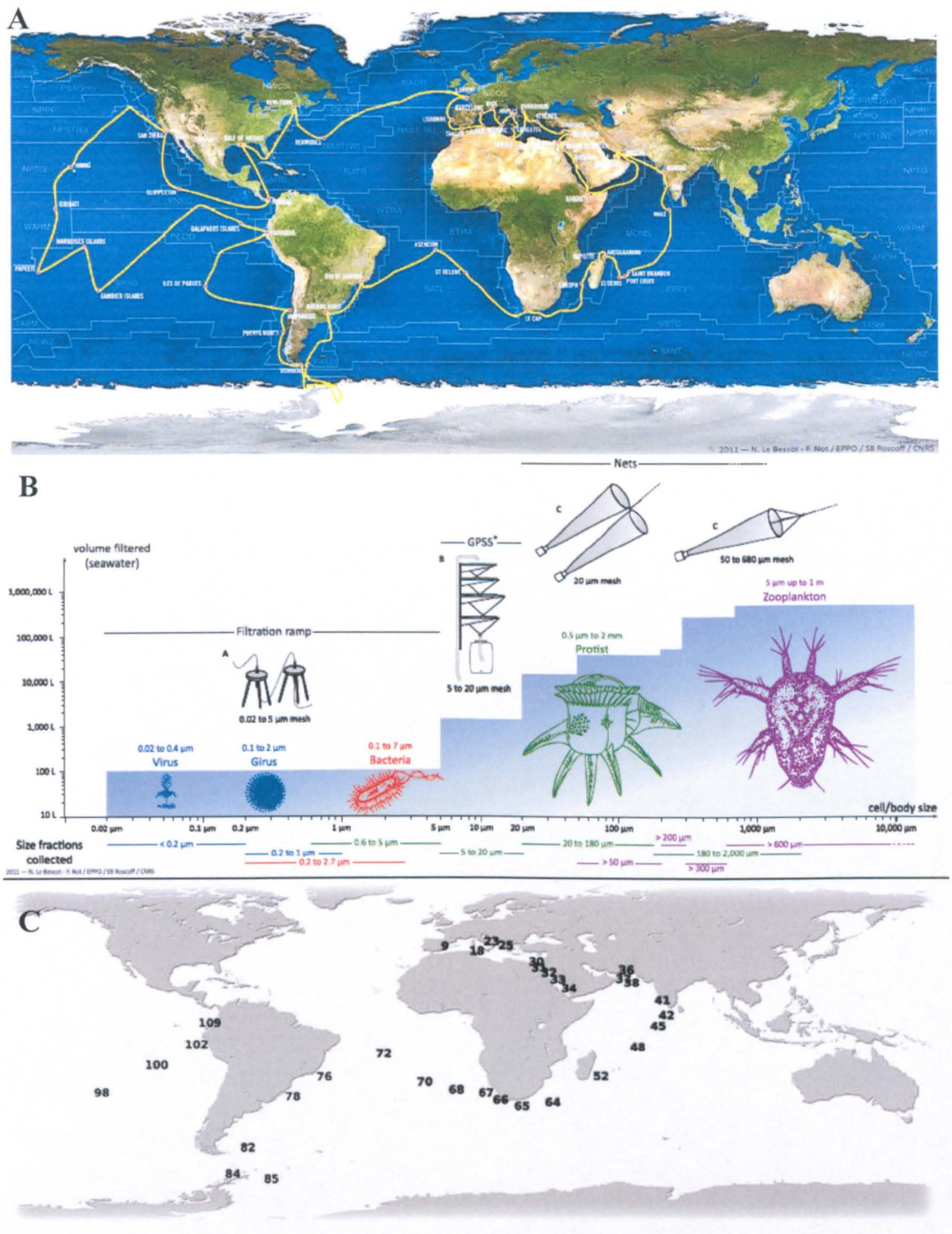


Fig. 3.1. *Tara* Oceans cruise. (A) Route of the *Tara* Oceans expedition. (B) Methods for sampling organisms by size classes and abundance. The blue background indicates the filtered volume required to obtain sufficient organism numbers for analysis. (C) Sampling stations for which samples were sequenced, source: (Karsenti *et al.* 2011).

the following parameters were set for the sequence retrieval procedure: i) do not compare both strands (i.e., do not include the reverse complement); ii) use global sequence identity; iii) cluster sequence to the best cluster that meets the threshold; iv) and a band-width of 20. Alignment coverage parameters and length coverage parameters were set to default settings. Only the clusters of those sequences that grouped most closely with the reference sequences of the *Leptocylindrus* and *Tenuicylindrus* species were selected for phylogenetic analysis. Sequences were annotated, i.e., given names, based on their smallest distance to a reference sequence (here on referred to as distance criterion), not based on phylogenetic criteria.

The clusters of sequences that grouped most closely with the reference sequences of the *Leptocylindrus* and *Tenuicylindrus* species were re-clustered at a sequence identity cut-off (similarity) of 0.97; other parameters were set as in the previous step. The result of this step was a series of secondary clusters of sequences with similarity $\geq$0.97. Only one sequence per secondary cluster was taken randomly to represent the cluster in the subsequent alignment and phylogenetic analysis procedure. This condensation of the sequence data was applied to keep the size of resulting trees manageable.

3.2.3 Alignment of obtained sequences. Sequence alignment involved the following steps. First, the clusters of sequences that grouped most closely with the reference sequences of the *Leptocylindrus* and *Tenuicylindrus* species were merged into a single database. The reference sequences of *Leptocylindrus* and *Tenuicylindrus* species were included. Sequences of related diatom species and of *Bolidomonas* were included as outgroups. The gathered sequences were aligned with the ClustalW module of Bioedit v7.1.3 (http://www.mbio.ncsu.edu/bioedit) and verified by eyeball to correct any miss-alignments.

3.2.4 Phylogenetic analysis of obtained sequences. Phylogenetic signal among parsimony-informative sites was assessed by comparing the measure of skewedness in the length-distribution among 100,000 random trees of the distinct sequences (G_1-value, PAUP*) with the empirical threshold values for 4-state characters, given the number of distinct sequences and parsimony informative sites (see Hillis 1992). Neighbour joining (NJ) trees for individual species datasets and the combined dataset of the V4 region and of the V9 region were constructed utilizing PAUP*. Trees were rooted with the most distant outgroup sequences. Metadata associated to the sequences (geographic location, sample date) were mapped over the resulting trees. Bootstrap support values were generated using 1000 bootstrap-replicated datasets and the same settings as in the NJ-settings.

3.3 Results

The results of the database searches have been presented separately because the BioMarKs dataset uses the variable V4 region, and the *Tara* Oceans dataset, the V9 region. The two regions are situated in different parts of the SSU rDNA.

3.3.1 *Pairwise dissimilarity.* Pairwise dissimilarities among the V4 sequences as well as among the V9 sequences of SSU rDNA of all of the *Leptocylindrus spp.*, *T. belgicus* and *Bolidomonas* spp. are presented in Table 3.2 and V9-region is presented in Table 3.3 respectively. Pairwise dissimilarities among the V4 region *Leptocylindrus* spp. varied from 0.00258–0.2066, between *Leptocylindrus* spp. and *T. belgicus*, 0.1749–0.2168. Pairwise dissimilarities among the V9 region ranged between 0.00 and 0.14 among *Leptocylindrus* spp. and from 0.16 to 0.20 between *Leptocylindrus* species and *T. belgicus*. Pairwise dissimilarities between the in- and outgroups varied from 0.15 to 0.23 for the V4 region and between 0.09 and 0.22 for the V9 region. A cut-off sequence identity value of 0.9 (dissimilarity 0.10) was chosen for the subsequent analyses.

3.3.2 The BioMarKs data (V4-region)

***i.* Sequence abundance.** *ggsearch36* search for *Leptocylindrus* and *Tenuicylindrus* sequences in the BioMarKs V4 dataset with a sequence identity cut-off of 0.9 resulted in a total of 4498 unique sequences. CD-HIT-EST-2D clustering with a sequence identity cut-off of 0.9 with 134 V4 sequences of reference species gave 3748, 5,116, 212, 60 and 56 unique sequences closest to *L. aporus, L. convexus, L. danicus, L. hargravesii, L. minimus,* and *T. belgicus,* respectively (Table 3.4, A). The percentage of similarity of the sequences to reference species is represented in Fig. 3.2, A. The sequences assigned to *L. aporus* were the most abundant and those assigned to *L. convexus* were the least abundant in the samples. Further clustering at identity cut-off of 0.97 for individual species resulted in 10, 1, 6, 21, 3 and 8 secondary clusters of multiple sequences assigned to *L. aporus, L. convexus, L. danicus, L. hargravesii, L. minimus,* and *T. belgicus,* respectively (Fig. 3.2, B). Clustering at other lower and higher values is presented in Fig. 3.2, B but the cluster generated at 0.97 was used for phylogenetic tree construction. The number of sequences in each secondary cluster is provided behind their end-nodes in the resulting trees.

ii. Alignment and phylogenetic inferences. The aligned dataset contained 91 distinct sequences and 419 characters (alignment positions) of which 149 were constant, 61 parsimony-uninformative and 209 parsimony-informative. Of the 91 sequences, two belonged to *Bolidomonas*, 28 to reference sequences of centric diatoms other than *Leptocylindrus* and *Tenuicylindrus*, and the remainder to sequences of those two genera. Evaluation of lengths of 100,000 random trees, given the sequence dataset, showed a skewed distribution with a G_1 value of -0.341252. The threshold value for 100 parsimony informative positions and far more than 25 different taxa is -0.12, which is closer to 0 than the observed value. Therefore, the V4 region contains significant phylogenetic information.

Table 3.2. Pairwise dissimilarities between V4 regions of the nuclear SSU rDNA of *Leptocylindrus*, *Tenuicylindrus belgicus*, *Bolidomonas spp.* and four centric diatoms

V4	1	2	3	4	5	6	7	8	9	10	11	12
1 *L. aporus*	-											
2 *L. convexus*	0.1653	-										
3 *L. danicus*	0.20403	0.19845	-									
4 *L. hargravesii*	0.20662	0.20103	0.00258	-								
5 *L. minimus*	0.16267	0.08247	0.18814	0.19072	-							
6 *T. belgicus*	0.2168	0.19327	0.20624	0.20886	0.1749	-						
7 *Bolidomonas mediterranea*	0.19419	0.17255	0.15939	0.15675	0.16189	0.18307	-					
8 *Bolidomonas pacifica*	0.19204	0.17025	0.18605	0.18341	0.16492	0.18081	0.08853	-				
9 *Paralia sulcata*	0.18529	0.17498	0.20112	0.20376	0.17502	0.15971	0.15053	0.14609	-			
10 *Stephanopyxis nipponica*	0.19482	0.1818	0.20273	0.20537	0.17648	0.16721	0.15262	0.15072	0.04691	-		
11 *Hyalodiscus sp.*	0.20239	0.21029	0.19985	0.20248	0.17902	0.19863	0.16298	0.17397	0.10678	0.10622	-	
12 *Melosira cf. octagona*	0.22284	0.20405	0.22773	0.23035	0.20394	0.15754	0.20974	0.19729	0.16578	0.15758	0.18115	-

Table 3.3. Pairwise dissimilarities between V9 regions of the nuclear SSU rDNA of *Leptocylindrus*, *Tenuicylindrus belgicus*, *Bolidomonas spp.* and four centric diatoms

V9		1	2	3	4	5	6	7	8	9	10	11	12
1	*L. aporus*	-											
2	*L. convexus*	0.08046	-										
3	*L. danicus*	0.10609	0.13503	-									
4	*L. hargravesii*	0.10609	0.13503	0.00000	-								
5	*L. minimus*	0.08621	0.08046	0.14080	0.14080	-							
6	*T. belgicus*	0.16523	0.20389	0.16192	0.16192	0.18169	-						
7	*Bolidomonas mediterranea*	0.17241	0.17241	0.17687	0.17687	0.22414	0.20476	-					
8	*Bolidomonas pacifica*	0.19540	0.16667	0.18288	0.18288	0.22414	0.21801	0.08046	-				
9	*Paralia sulcata*	0.13218	0.17241	0.14697	0.14697	0.18391	0.18139	0.14943	0.18391	-			
10	*Stephanopyxis nipponica*	0.10042	0.15917	0.12627	0.12627	0.16481	0.16285	0.16490	0.20040	0.11818	-		
11	*Hyalodiscus sp.*	0.09373	0.16373	0.14394	0.14394	0.16956	0.16692	0.14015	0.16992	0.12851	0.07632	-	
12	*Melosira cf. octagona*	0.12069	0.16092	0.14096	0.14096	0.15517	0.20985	0.20115	0.21264	0.13218	0.12357	0.14030	-

The NJ tree inferred from the V4 dataset of BioMarKs is presented in Fig. 3.3. With *Bolidomonas* sequences designed as outgroup, the sequences assigned to *Leptocylindrus* grouped into a basal grade consisting of two clades, one of which contains the reference sequences of *L. hargravesii* and *L. danicus* and the other one the reference sequences of *L. convexus, L. minimus* and *L. aporus*. Next to branch off was a clade containing all sequences of *T. belgicus.* All other (centric diatom) taxa included in this study grouped in a clade as sister to *T. belgicus*, though without bootstrap support. Moreover, the centric diatoms in this clade did not form clades as expected according to their taxonomic assignment, i.e., bi- and multi-polar centric diatoms did not form a clade.

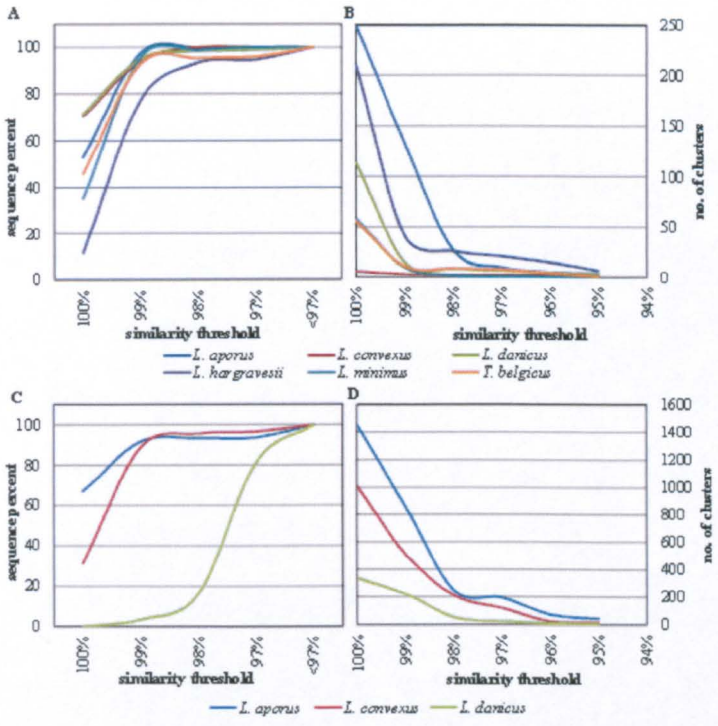


Fig 3.2. Percent of sequences that form a cluster with the reference sequence (A and C) and the number of OTUs detected at different OTU calling thresholds for the datasets (B and D) retrieved from environmental samples. (A) the percent of sequences that form a cluster with the reference sequence at different OTU calling thresholds based on V4 for the BioMarKs dataset. (B) the number of OTUs detected at different OTU calling thresholds based on V4 for the BioMarKs dataset. (C) the percent of sequences that form a cluster with the reference sequence at different OTU calling thresholds based on V9 for the *Tara* Oceans dataset. . (D) the number of OTUs detected at different OTU calling thresholds based on V9 for the *Tara* Oceans dataset.

Table 3.4. Summary of number of sequences retrieved from the environmental sequences database. (A) BioMarks. (B) *Tara* Oceans

A

Species	Unique	Total abundance	Clusters	No. of stations
L. aporus	3748	48877	10	4
L. convexus	5	17	1	1
L. danicus	116	692	7	2
L. hargravesii	212	671	21	2
L. minimus	60	248	1	1
T. belgicus	56	172	8	2

B

Species	Unique	Total abundance	Clusters	No. of Stations
L. aporus	1456	41700	147	32
L. convexus	1012	31265	85	31
L. danicus	338	4666	26	20

Six clades (marked I - VI) were recognised for *Leptocylindrus* species and *T. belgicus*. Each clade, except one, contained multiple end-nodes and several of these end-nodes represented secondary clusters of multiple sequences with similarity ≥0.97 (see Fig. 3.3). For each of the clades, the number of end-nodes, the total number of sequences, the reference species, if any, and the geographical origin of its sequences, has been presented in Table 3.5 and Table 3.7. The number of sequences in each secondary cluster is shown behind the end-nodes in Fig. 3.3.

Clade I (100% bootstrap support) included the reference sequences of both *L. danicus* and *L. hargravesii,* 27 end-nodes of BioMarKs sequences / secondary clusters, as well as four environmental sequences from GenBank. The reference sequences of *L. danicus* and *L. hargravesii* can be distinguished by only two base changes in the V4 regions and were recovered closely together well inside this clade. Notably, the reference sequence of *L. danicus* grouped with 685 sequences in one large secondary cluster and the reference sequence of *L. hargravesii* grouped with 635 sequences in another large secondary cluster, whereas all the remaining sequences in Clade I formed end-nodes containing singletons or

pairs. The marked sequence variation among all these remaining sequences did not result in any significant bootstrap support for clades.

Clade II (100% bootstrap support) included the reference sequence of *L. convexus*, which was identical to a secondary cluster of 17 sequences, and three GenBank sequences. All these sequences were found to be closely related.

Clade III (100% bootstrap support) contained the reference sequence of *L. minimus,* with four secondary clusters of sequences, each with many sequences. The reference sequence was recovered branching off first within this clade.

Table 3.5. Summary of BioMarKs NJ trees. The number of end-nodes within each clade, their total sequence abundance and the geographical origin of the sequences are presented.

Clade	No. of end-nodes	Sequence abundance	Reference Species	Geographical location
Clade I	33	1363	*L. danicus /hargravesii*	Naples, Oslo
Clade II	5	17	*L. convexus*	Naples
Clade III	5	246	*L. minimus*	Oslo
Clade IV	1	1	-	Naples
Clade V	9	48876	*L. aporus*	Naples, Oslo, Barcelona, Gijon
Clade VI	172	8	*T. belgicus*	Naples, Oslo

Clade IV was found to contain a singleton only; it contained no reference sequence.

Clade V (100% bootstrap support) contained the reference sequence of *L. aporus,* as well as eight additional end-nodes, four of which consist of secondary clusters. The reference sequence itself groups in a secondary cluster with 48820 additional sequences. The clade showed marked sequence variation, but the reference sequence was recovered well inside it.

Clade VI (100% bootstrap support) contained the reference sequence of *T. belgicus,* as well as eight additional end-nodes, one of which was found to consist of a secondary reference sequence. The remaining singletons were found to be remarkably different, but without any phylogenetic structure.

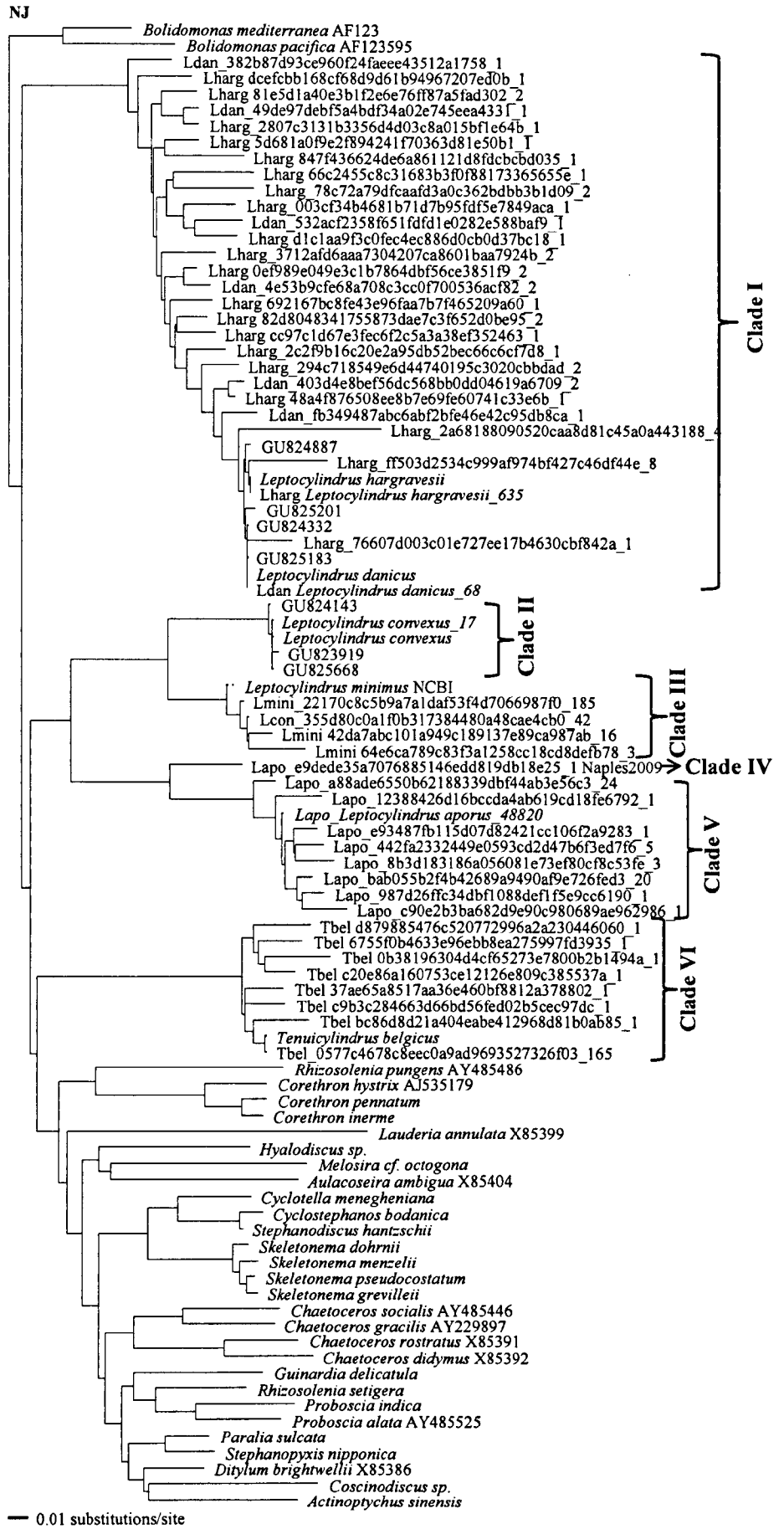

NJ
Bolidomonas mediterranea AF123
Bolidomonas pacifica AF123595
Ldan_382b87d93ce960f24faeee43512a1758_1
Lharg dcefcbb168cf68d9d61b94967207ed0b_1
Lharg 81e5d1a40e3b1f2e6e76ff87a5fad302_2
Ldan_49de97debf5a4bdf34a02e745eea433f_1
Lharg_2807c3131b3356d4d03c8a015bf1e64b_1
Lharg 5d681a0f9e2f894241f70363d81e50b1_1
Lharg 847f436624de6a861121d8fdcbcbd035_1
Lharg 66c2455c8c31683b3f0f88173365655e_1
Lharg_78c72a79dfcaafd3a0c362bdbb3b1d09_2
Lharg_003cf34b4681b71d7b95fdf5e7849aca_1
Ldan_532acf2358f651fdfd1e0282e588baf9_1
Lharg d1c1aa9f3c0fec4ec886d0cb0d37bc18_1
Lharg_3712afd6aaa7304207ca8601baa7924b_2
Lharg 0ef989e049e3c1b7864dbf56ce3851f9_2
Ldan_4e53b9cfe68a708c3cc0f700536acf82_2
Lharg 692167bc8fe43e96faa7b7f465209a60_1
Lharg 82d8048341755873dae7c3f652d0be95_2
Lharg cc97c1d67e3fec6f2c5a3a38ef352463_1
Lharg_2c2f9b16c20e2a95db52bec66c6cf7d8_1
Lharg_294c718549e6d44740195c3020cbbdad_2
Ldan_403d4e8bef56dc568bb0dd04619a6709_2
Lharg 48a4f876508ee8b7e69fe60741c33e6b_1
Ldan_fb349487abc6abf2bfe46e42c95db8ca_1
Lharg_2a68188090520caa8d81c45a0a443188_4
GU824887
Lharg_ff503d2534c999af974bf427c46df44e_8
Leptocylindrus hargravesii
Lharg Leptocylindrus hargravesii_635
GU825201
GU824332
Lharg_76607d003c01e727ee17b4630cbf842a_1
GU825183
Leptocylindrus danicus
Ldan Leptocylindrus danicus_68
GU824143
Leptocylindrus convexus_17
Leptocylindrus convexus
GU823919
GU825668
Leptocylindrus minimus NCBI
Lmini_22170c8c5b9a7a1daf53f4d7066987f0_185
Lcon_355d80c0a1f0b317384480a48cae4cb0_42
Lmini 42da7abc101a949c189137e89ca987ab_16
Lmini 64e6ca789c83f3a1258cc18cd8defb78_3
Lapo_e9dede35a7076885146edd819db18e25_1 Naples2009
Lapo_a88ade6550b62188339dbf44ab3e56c3_24
Lapo_12388426d16bccda4ab619cd18fe6792_1
Lapo_Leptocylindrus aporus_48820
Lapo_e93487fb115d07d82421cc106f2a9283_1
Lapo_442fa2332449e0593cd2d47b6f3ed7f6_5
Lapo_8b3d183186a056081e73ef80cf8c53fe_3
Lapo_bab055b2f4b42689a9490af9e726fed3_20
Lapo_987d26ffc34dbf1088def1f5e9cc6190_1
Lapo_c90e2b3ba682d9e90c980689ae962986_1
Tbel d879885476c520772996a2a230446060_1
Tbel 6755f0b4633e96ebb8ea275997fd3935_1
Tbel 0b38196304d4cf65273e7800b2b1494a_1
Tbel c20e86a160753ce12126e809c385537a_1
Tbel 37ae65a8517aa36e460bf8812a378802_1
Tbel c9b3c284663d66bd56fed02b5cec97dc_1
Tbel bc86d8d21a404eabe412968d81b0ab85_1
Tenuicylindrus belgicus
Tbel_0577c4678c8eec0a9ad9693527326f03_165
Rhizosolenia pungens AY485486
Corethron hystrix AJ535179
Corethron pennatum
Corethron inerme
Lauderia annulata X85399
Hyalodiscus sp.
Melosira cf. octogona
Aulacoseira ambigua X85404
Cyclotella menegheniana
Cyclostephanos bodanica
Stephanodiscus hantzschii
Skeletonema dohrnii
Skeletonema menzelii
Skeletonema pseudocostatum
Skeletonema grevilleii
Chaetoceros socialis AY485446
Chaetoceros gracilis AY229897
Chaetoceros rostratus X85391
Chaetoceros didymus X85392
Guinardia delicatula
Rhizosolenia setigera
Proboscia indica
Proboscia alata AY485525
Paralia sulcata
Stephanopyxis nipponica
Ditylum brightwellii X85386
Coscinodiscus sp.
Actinoptychus sinensis
Clade I
Clade II
Clade III
Clade IV
Clade V
Clade VI
0.01 substitutions/site

Fig. 3.3. Neighbour joining tree illustrating the relationship among the V4-based secondary clusters from the BioMarKs dataset along with *Leptocylindrus* species, *Tenuicylindrus belgicus* and other diatom species. Bootstrap values have been generated with 1000 replicates. Bootstrap support for clades receiving <50% support have been omitted.

iii. Phylogeography and additional metadata: With eight sampling sessions at six sample stations, sequences attributable to the genus *Leptocylindrus* and *T. belgicus* were found at four stations, namely, Oslo, Gijon, Barcelona and Naples. The remaining two stations (Roscoff and Varna) showed no sequences of *Leptocylindrus* and *T. belgicus*. The relative sequence abundance of the *Leptocylindrus* species and *T. belgicus* during sampling of 2009 and 2010 have been shown in Fig. 3.4. Sequences assignable to *L. danicus* and *L. hargravesii* have been pooled in the geographic analysis because they grouped in a single clade in Fig. 3.3.

The samples taken at the Oslo Fjord-station in September showed the presence of *L. aporus, L. danicus / hargravesii, L. minimus* and *T. belgicus*, whereas in July 2010, only *L. danicus / hargravesii* and *L. minimus* sequences were present, of which the latter were the most abundant. Notably, *L. convexus* was absent in both samples from Oslo. The sample taken at the Gijon-station in September 2010 contained a few sequences of *L. aporus*, and so did the sample taken at Barcelona in February 2010. All other species sequences were absent from these samples. In the coastal Mediterranean station off Naples (LTER-MC), sequences of *L. aporus, L. danicus / hargravesii* and *T. belgicus* were obtained during 2009 autumn sampling and those of *L. aporus, L. convexus* and *L. danicus* during the 2010 spring sampling.

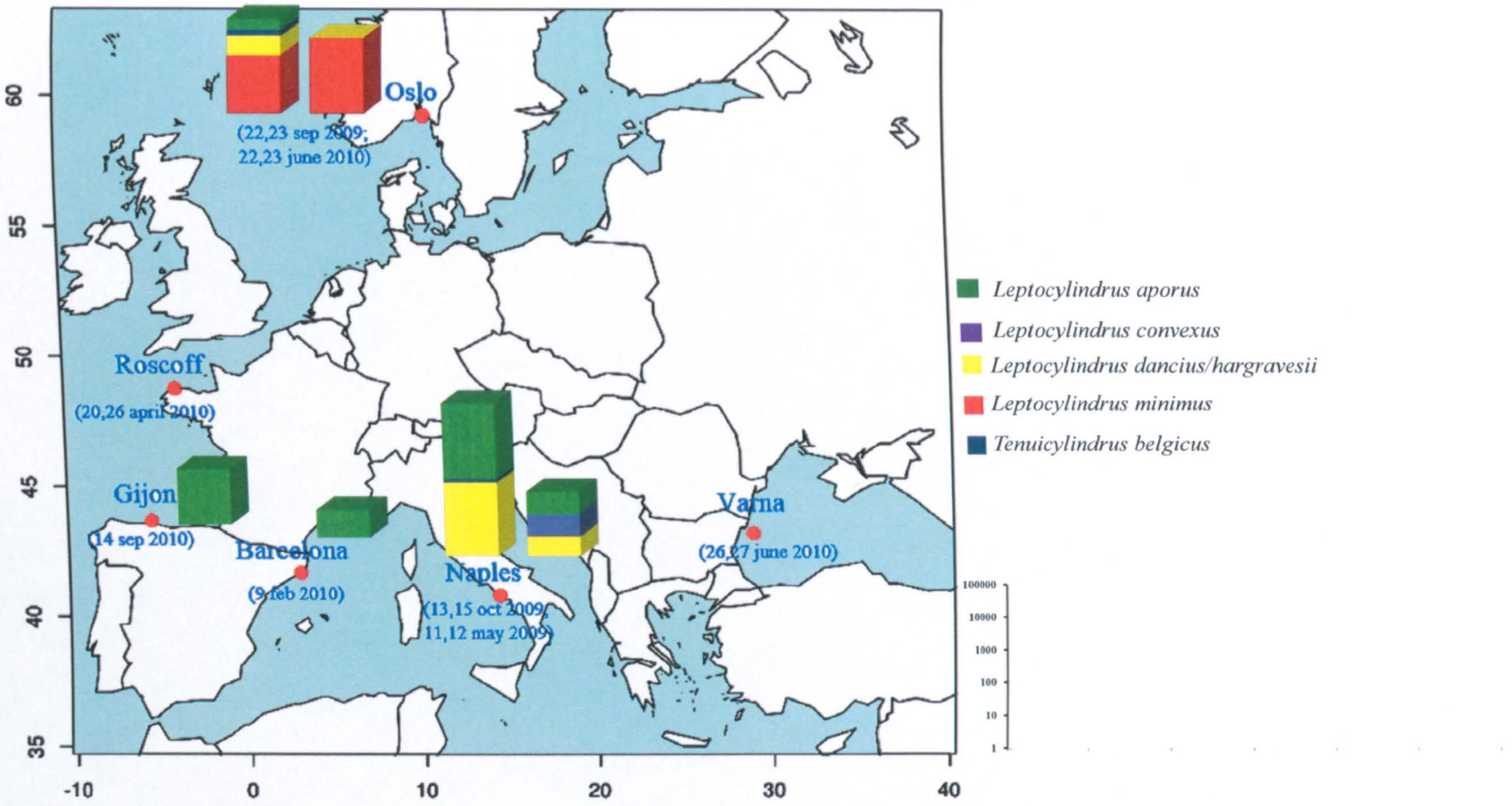


Fig 3.4. Distribution and sequence abundance of *Leptocylindrus* species and *Tenuicylindrus belgicus* in the samples of the EU-BioMarKs dataset. Sequences abundance is presented in $\log_{10}$ values.

Sequences belonging to *Leptocylindrus aporus* and *L. danicus / hargravesii* were detected in samples taken in seawater with temperatures ranging between 12.5 and 22.8 °C (Fig 3.4, Table 3.5). However, both were absent in the June 2010 water column samples from Oslo when the temperature was within this range. Contrastingly, *T. belgicus* was present in the October 2009 (22.8°C) samples from Naples and September 2009 (15.5°C) samples of Oslo. *Leptocylindrus convexus* was found only in the May 2010 sample from Naples when water temperature was 13.5°C.

Sediment samples collected to trap out of season species that form resting stages showed the same sequences of species recovered in subsurface and DCM. The only exception was that of *L. danicus / L. hargravesii,* whose sequences during June 2010 in Oslo were found exclusively from sediments.

3.3.3 *Tara* Oceans data (V9-region)

i. Sequence abundance. CD-HIT-EST-2D clustering of the *Tara Oceans* V9 dataset of Bacillariophyta at a sequence identity cut-off of 0.9 with 133 sequences of reference diatom species resulted in total of 2806 unique *Leptocylindrus* sequences. (Unfortunately, the sequences of *Tenuicylindrus* were excluded because the species previously unknown to science and given its phylogenetic position was considered outside Bacillariophyta). Of these 2806 sequences, 1456, 1012 and 338 belonged to *L. aporus, L. convexus and L. danicus* respectively; the percent similarity with reference sequences is presented in Fig. 3.2, C. The total abundance of these sequences is presented in Table 3.4, B and their percent similarity distribution with the reference sequences is represented in Fig. 3.2, D. Sequences assigned by the distance criterion to *L. aporus* were the most abundant and those assigned to *L. danicus / L. hargravesii* were the least abundant. Further clustering with CD-HIT-EST at a percent identity of 0.97 for individual species resulted in 147, 85 and 26 clusters for *L. aporus, L. convexus, and L. danicus / L. hargravesii,* respectively

(Fig. 3.2, C). Clustering at other lower and higher values is presented in Fig. 3.2, D but the cluster generated at 0.97 was used for phylogenetic tree construction. The number of sequences in each secondary cluster is provided behind their end-nodes in the resulting trees.

ii. Alignment and phylogenetic inferences. The dataset contained 291 distinct sequences and 147 characters of which 26 were constant, 17 were parsimony uninformative and 104 were parsimony informative. Of these 291 sequences, five belonged to non-diatom photoautotrophic stramenopiles, 28 to centric diatoms outside *Leptocylindrus*, and the remainder to sequences assigned to this genus. Evaluation of lengths of 100,000 random trees, given the sequence dataset, showed a skewed distribution with a G_1 value of -0.1936. The threshold value for 100 parsimony informative positions and far more than 25 different taxa is -0.12, which is closer to zero than the observed value. Therefore, the V9 region contains significant phylogenetic information.

The distance tree inferred from the V9 dataset of *Tara* Oceans is presented in Fig. 3.5. In this tree, eleven distinct clades (Clade I to Clade XI) were recognised. All these Clades contained two or more end nodes, many of which constituted secondary clusters of sequences less than 3% different. The numbers of end-nodes in each clade, the total sequence abundance, the reference species if any and the geographical locations in which the sequences have been sampled are presented in Table 3.6. Bootstrap support for most of the clades was found to be insufficient (clades supported are indicated).

Clade I was shown to contain 11 end nodes of which six are singletons; one of the secondary clusters contained 261 sequences. The clade grouped closely with radial centrics *Aulacoseira*, and *Actinoptychus* and the bipolar centric diatom *Chaetoceros.*

Clade II contained the reference sequence of *L. aporus*, and 49 end-nodes, 37 of which were singletons. An end-node close to the reference sequence consisted of a secondary

NJ

Bolidomonas mediterranea AF123596
Bolidomonas pacifica AF123595
Phaeothamnion confervicola AB365203
Spatoglossum pacificum AB087132
Dictyota dichotoma AB087107
Coscinodiscus sp. GGM-2004 AY485448
Tenuicylindrus belgicus
Stephanopyxis nipponica M87330
Chaetoceros didymus X85392
Porosira pseudodenticulata X85398
Proboscia alata AY485525
Attheya cf. septentrionalis AY485517
Attheya longicornis AY485450
Skeletonema menzelii AJ632217
Skeletonema grevilleii CCMP1685 DQ396512
Skeletonema dohrnii B104 AJ632210
Ditylum brightwellii X85386
Skeletonema pseudocostatum X85393
Cyclotella menegheniana WC03-01 DQ514861
Corethron pennatum
Corethron inerme
Thalassiosira rotula X85397
Stephanodiscus hantzschii DQ514914
Stephanodiscus niagarae DQ514908
Chaetoceros rostratus X85391
Guinardia delicatula AJ535192
Corethron hystrix AJ535179
Rhizosolenia setigera M87329
Rhizosolenia pungens AY485486
lapo_c564feec5c3f6a6494fece7d9ffd77e1_1
lapo_13a285544b71de3c5f7f132f72b94d98_2
lapo_8e43eba3f7a4ca8d03d7e6a88f6f6504_1
lapo_29742085807b380bdfdbc9a8359dae49_13
lapo_9a416003953352c283edde7416445fb4_1
Actinoptychus sinensis p394 AJ535182
Aulacoseira ambigua X85404
Chaetoceros socialis AY485446
Chaetoceros gracilis AY229897
lapo_e4ae9d0e3af91ac48fd1fc4be5db0639_1
lapo_af367ff5cf256f82156ad3075b192091_1
lapo_ee3cab6534c55272f3a0092f3ebba279_2
lapo_34a2938d44b226246bb6a43538a12438_7
lapo_8a4f47801d0e82ee058e59d42853b523_2
lapo_03285e428d1a456ee31602ae14d16217_**261**
lapo_cc2cc75bb435199de30cd69047a1be7e_1
lapo_c80c5658162214e1f4d3e5d88686e569_1
lapo_27005d9ffc2b2712f1427565f0bbe8b5_9
lapo_365e8158587344bf5f936b19d83b725b_1
lapo_4e5a045c9a56ee5eb7812a8e7b323283_1
lapo_0a5f1e91172a1a41cc19614333788ed3_1
lapo_2cdbe753457a60da113bc43bce42d2cd_1
lapo_103ae6cdd9955a8a8ee91caeb070e43b_1
lapo_2361a02179b10a5dc00c771326ad643f_1
lapo_0c0b0ae8b04c371ad0f6ea5c57a0a989_1
lapo_ceed6c1f4931f420fc134c040b2810ce_2
lapo_71fcd5a3aecc562dc522bacb69ca24b4_1
lapo_f9659d8dfd2ae4e1f00a51780df83dac_1
lapo_1410f3d373d6a687ab3a9e0307dd9a8a_2
lapo_5d4413b9fa175fd432dc9a9c06f03bcc_1
lapo_33bd3edeabbbc80704d922745c96a7d3_1
lapo_1ce43a71a6cc03dda625063afff7adfa_5
lapo_51d55ee7dde47542c82c45e85353fccd_1
lapo_ae97363f9e126773356169713b0b359b_1
lapo_3139e3d4d7e9b32177a5afad88d84c1c_1
lapo_b7dbaf6b67b82e0d28318d85aa7dada5_1
lapo_51b76c308747aebbe9aaa7268f5d36ab_1
lapo_815f7ed4115429fdf259e59a1db8c6ee_1
lapo_234ae4d45e9f11b72b386fc32876fecb_2
lapo_21860bfd00cd649c68342049aa20a9dc_1
lapo_371cb358f41d6e4bc12699e270ae4af0_1
lapo_298725568907bcc110f5abc83f0bf90c_2
lapo_1244f29bd0c1448fcfbd4d944832e975_5
lapo_b7057e903285f9265c341b5932f61e7d_13
lapo_83295cd5bc9ecc13f1fcdae321a5de56_1
lapo_5ee474a923b1f5681d714ae573fbb974_2
lapo_ae4879faac9aa146ffb7b3cf457b1fd4_1
lapo_98ffe5a5396e77514f01e4de544d8134_1
lapo_2f57aa61cfdd02ffc7368952aa48eb2d_**20**
lapo_6adbff9d112a9fe8c198c76ffce10391_1
lapo_4a9d8a67b05ad0dc2164e7b69ac20f5e_1
lapo_a4e49628f5e0f9f0c7e997d79adb5d10_1
lapo_562e08903953e601191cddcd1eaec5e8_6
lapo_04d1c235f5ac1b74c8e716e3e07b1fbe_**24**
lapo_ca8831f3d945fc9f03553b0f4271f9cc_1
lapo_1915962752c5d11b9a660af9287f3117_1
lapo_2d7a940e9febb4d8e702b1538f8e2689_1
lapo_7ef9fd22fb858bd9ad8e75bfa67dfb95_1
Leptocylindrus aporus
lapo_000f2d490b0bc7c615182ec07fe0b186_**38943**
lapo_3406e9c2402df1402dd648551e257fe0_**25**
lapo_25ea2f5e52cf6a083764215d029584cb_1
lapo_36edf91dcf93dda3a500b445a802d8a6_1
lapo_e0dda7c5d40cf2acd39930cc999f3984_1
lapo_f4239460d9e66951b2f1ea012a34807f_1
lapo_798ed45a641ef0c154608b4727d45f9b_1
lapo_790a289a734cc4b460e8d1a221420376_1
lapo_9f56a691cbc22db0e2f4ee2ae07757be_1
lapo_b98f279e293906ba98171d908e31cd64_1
lapo_c53b4480d07a6dc24400613c4cd59f07_1
lapo_c71755ed64960d5edc523e67b7e4153e_15
lapo_7b9f025afa24b1a10ff0a8041bd27130_**33**
lapo_06c60896f93e51063ebfb493f3719510_**35**
lapo_097c36f3c61700a426b4cb4d40e52fa8_15
lapo_63691802d1953a6ec1a5d2b414446bf9_1
lapo_33d61fdbf52ea1e7b535bf69bad0f75f_5
lapo_489d94a38f05fa8b5c4a5659a9454c97_2
lapo_295ec4fca8304b64b2ea6c2de909751f_**23**
lapo_00f271b46b567b8d26235e1db25698a3_**21**
lapo_b47e52bfaab55674f145c6fd11ebbaa8_1
lapo_2fd088d33f8e939ec4dab6f3408a78e5_2
lapo_51b41e8a0239939b047d0e8a5378034f_1
lapo_34fc95c2e40769d13acf2c8231c46beb_**28**
lapo_b768fdb0cac97ded543405a92f9b9a02_1
lapo_3f54344a0bd7c8ded34d8195bef944a0_5
Ldan_98169d911766b56a990b90a957a885db_1
Ldan_5104042e25b67d997d696aad32440629_1
Ldan_0ca9e9598072f125309ed4c92a14596d_**409**
Ldan_2ed9061668f02a8adf11a20a4587649c_1
Ldan_5c35d4c65973220d253fc4fa1b7bee79_1
Ldan_d7a15bdcd12441f0035abcd69327904b_1
Ldan_1e26b9d13a930a187129d56ef9bacfa3_**132**
Ldan_cad86de0e9d825f1a4d9b1cef87109d3_1
Ldan_*Leptocylindrus_danicus*_**180**
Leptocylindrus danicus
Leptocylindrus hargravesii
Ldan_c024b9bc6d297512d243c3d73604114c_1
Ldan_c5b4de79440bd4cbe7366a1e2998c905_2
Ldan_a35c3013dc8827c444df31c0a346e1fa_1
Ldan_2a2363a830012a8a377fce7fa5126f7f_3
Ldan_5db11bb2b48837ed9859d108943b4872_**20**
Ldan_d41e5c7a7c3a016d76a5ea7042da7949_1
Ldan_61cfca380fca881cbefc376e87c78d52_1
Ldan_20ce3aabcc5e9de19875c5b8c28c53b4_13
Ldan_4f7f9915e08a3c7250cea7f2e01a31b6_4
Ldan_10f4d98eb7e1d179fca1122e709d4e07_5
Ldan_a7434515c30d630d888443e6c6a06ea4_1
Ldan_abfa277c60337473eda729090f877c52_1
Ldan_04d3c12cabd001970a225379ebf40177_1
Ldan_6a114a3ff42c33b1779b129d4a897ea4_1
Ldan_003ed200f87b52a01207361444e49fdb_**3252**
Ldan_0c41e230df4584e17abb0d792f223bab_1
Ldan_6e0f91cbf1c1373bc639854778ad3bbd_1
lapo_095066b9f1baf6e8007b34cbcb1e36f5_1
lapo_562d143d141bead0579d1957b8f5fc7f_2
lapo_147f3d1ada92370cef84a08da81cb3b1_**23**
lapo_29bc71f8025f71af11fee2096bb1a39f_19
lapo_3ce7700e8929ed0f2e054e58dedc9b05_2
lapo_0d739d1112ffdaa7da87e2e2cbccbb6e_16
lapo_039d0d2b72a058a13582b44648612ec6_**134**
lapo_24e6e171d8500eea1a3002fd080ea0ca_2
lapo_50a16a57402f6ab5756e465890aa8720_1
lapo_671c7509d910beccbe9ea37619ac8ab9_**37**
lapo_0b369d2f41e11c5064526ba2d0b1c5f1_**701**
lapo_a919053ce68290486de9cc58511f96f7_1
lapo_d00dc0c584b07bb40e1181fc3a88b4c7_1
lapo_b95c838da06c3b529b75007d377921c9_1
lapo_a30cd0b698e6275911fd98f8cb8579a2_1
lapo_e828c8bddd772cf1e33b9c9ab5d91ef7_1
lapo_f0077b32cb53e472590fe6eac1eddd05_1
lapo_f7c76b1fc67ceea9a34d3b5d03c668b3_1
lapo_fe75251e651460e1ed68a002b057d484_1
lapo_328a3e4d8dc8d90a9b0d8477e0948912_2
lapo_61cc73f418f902f702480c637c69ec07_1
lapo_bdf8c568a11f9ef1f4cde4054aabd176_1
lapo_ce1bee9c08cb29b1cc52118b7e9dea7d_1
lapo_9fe6ef4c9c261bce2c5aeffffaec4906_2
lapo_1add05836047480fb4c95d4de502a48b_12
lapo_d34c925337cf97280586fcf4b28625df_1
lapo_0eb3a3a0f01378cd65a097527790e0c0_**21**
lapo_f963c88f6db629167403b583209c511e_2
lapo_17790f6aee13de80247e8b7bb6184510_1
lapo_985e8bc3f14262f5af476090012d1f5a_1
lapo_ae82e0ef9f31f0988bfd3fe438be327f_1
lapo_9867d15e870d487efc5049c568d95b44_1
lapo_c27ccb517f5355071c815fe7f2cc1a20_2
lapo_103d1e7ae4ab2b069ae8462ddaac6b1a_**245**
lapo_0275adea3cb7370e7d7a806f2c4edd2e_**144**
lapo_0da1bd2bec82e83ed7ec523ea3ae830c_**201**
lapo_21c8ecb03eaac886559e6995da6527b7_2
lapo_58120feb41ae44697220c01aaefc981f_1
lapo_c2ac2dbb18c867da6c6a96158e341cc4_10
lapo_5c1e2eaa45b1132483512e6cce9b0c6c_5
lapo_ef8a90c32f4beeb8a56969cec305299e_1
lapo_0228729ef43f583af60fc3b99db10837_**27**
lapo_f3f829c2b89ccd992d6280928124d025_1
lapo_15b28c50148d9ffa264a565da6cb02ce_**21**
lapo_043d62780d89cd0c2a52137106072acc_1
lapo_8469f243a232f52c7af0bbecd9fb2633_3
lapo_14fe46e08b0d39da695fff55b741a559_**27**
lapo_52da24d61d7e7509688dfdf5eb54d6bf_1
lapo_00e6651a351ecb10943bbc29b0924731_**24**
lapo_c86091460c1980aa1d8b8aa66c34712d_1
lapo_08ba0e2eb91ba291094be0443e01fb55_**55**
lapo_2baeb71061f8acc2692fc5604e18a601_1
lapo_3fd5dae1e805fdf73f05d9264e827c3a_7
lapo_b1d337c62f429d87408ab643425a2105_1
lapo_d1c02b0e470e7d3702a0d788a7094377_1
lapo_741bdaef8d6f17d7b4ef5cfa6fcc5896_1
lapo_187b54dfd9b1bd43563bc2c3accdbb3a_1
lapo_3020e3888a9822b2ac7f294a2b3c9862_**48**
lapo_3e463d3bc6b1e95bc9f7d4afed9b2d3f_3
lapo_92f47f8ed0993fa000d0b35eb1b6bc36_1
lapo_7505ded1f9c38977cfd8baeb57f76e34_1
lapo_a227b55272448a43ef4c767545afda14_2
lapo_01b7025410c0c16cd25daa798ddffd89_**283**
lapo_5215b61ba60689dcc3b52ed6e01d7eeb_1
lapo_49fe5295b27db875d98c997489f359ca_1
lapo_6222eac4319656f7f80dbd94b278b7b4_1
lcon_0a48bdf22b29f8be49368943e3af9401_2
lcon_1f23e1b0de41bb56224da943a21445fb_6
lcon_f958002a50926f532dc5ba69a89d9435_1
lcon_e6093e5193450b867d7d54bc5dd3a2b8_1
lcon_0211ce3af39d2d72f5eda6c5c38998e1_**625**
lcon_312098305b66007449a378c2b9c8bc1d_**27**
lcon_3cff101a2d45818f6af30dfe585f2c80_2
lcon_1f4a56169351dc2df0ecb97f88ce95f4_1
lcon_2704077db0f6147b28a465108ffbe73f_2
lcon_165d52d7cec09fc100fc1f517127f606_**117**
lcon_39791f53bf7637d42d999ed9ef71a078_1
Leptocylindrus minimus
lcon_231214cd742f5d37be317e81a104e616_1
lcon_4a2b36bc7eb8e646a04139059a28e0cb_1
lcon_2a864f5f7abf19e620a8524a553d442d_**77**
lcon_30191d18b824dfc3369a5b9023205ff5_1
lcon_30ed29056c95d8389278449ca236990a_3
lcon_6f01582e37de500c3212efef6903dc60_2
lcon_fa3e61f1a9c17c3ccb17b2ddfb9d1751_1
lcon_22a8248d470c929a3a349911873e5887_1
lcon_ecc7172e35f0419e84358e6b704fd91e_1
lcon_730a75eabf0e8a7071d15bdcc04b790e_1
lcon_a81b4ee954e73fc24c20bb54413fe2d6_2
lcon_d34b95d090754abbf001d10b8e83dd35_1
lcon_34883bbe2cef3669f37c591e3c68638a_2
lcon_9a1cbfb12dfd06a4a95f7d00553258b7_1
lcon_749e0a2d0cca3b258278798ced604fa5_1
lcon_a5acf2ad7b6be57e7ddd409c55537302_1
lcon_542ff16cc574d85edb80e8ebf87af151_2
lcon_d222af4aeb38f6eff3e116d64d880a8b_1
lcon_1ffdc6a7bc2801ecefdfc9827acd27dd_9
lcon_21bf647e00bf28150820c2983cf5e55f_1
lcon_5edb57d558484a4ec5c64f8f33641e73_2
lcon_2e914dc3c04a1950eeaab49a5adb7828_1
lcon_33871cf717478d79249dfe4389f7f5df_1
lcon_3f3e07c22b019d3fb2c5fc0a00445359_1
lcon_*Leptocylindrus_convexus*__**30135**
Leptocylindrus convexus
lcon_19c7d7da56cdb6651a7b424416a69687_1
lcon_d9dc5101daa044ad5eb0323daabd4934_1
lcon_0bae4e8f3120abcf99ba237d4f56dd10_5
lcon_57d172ebb7314d116dd013fcd56a781d_2
lcon_592618adece32760a90986e4fb04879a_1
lcon_c135570a4fad7cad61e119a850fdf194_1
lcon_4607a1c7b56bc13c3f987de67517e47e_1
lcon_fa3dc28d7e3b25690e93cec4553496da_1
lcon_554c40428175dc1d3a4c54c1373e1985_1
lcon_8b5c7ceded0fda7f4d6f91f039980df1_1
lcon_9776aa98a5e2f24a9ae8afeecf63bb26_1
lcon_0493efa84d60850232815182d62da711_1
lcon_240d4dfe0cfc3d0f9767812a59e60dcd_1
lcon_b8ac5063a39eb88ca59ff71c4472038e_1
lcon_2e44cd13af24148e7519bb57931e378e_19
lcon_ecf45c2ab4c7ded4103414acf60ffeeb_1
lcon_443b0460c32c027083cc3865d4201ae0_15
lcon_9b40dede775c5e8fcc158e8724060b78_2
lcon_dcf46968e2725477c627651b791df80c_1
lcon_0daf5f5593aabe0298321162d9b74239_19
lcon_e238d9e09a32943ec31e7ad69dcff67d_1
lcon_4bdd9ed75b08e48cbab86d4780bbc727_1
lcon_83513f8c7df47338c7cc9f3be7689690_1
lcon_c66e02d910d9badfa17e3ec67aa607d8_1
lcon_9329a5fd559b0618fd94bd6df668e6f4_1
lcon_747ebb7d258979633a6cbe6c832f542c_1
lcon_d0db38d655ec3ad83e06601557c6d1ec_1
lcon_26bd2d2c3014cfb4d50c0d964bff59fa_1
lcon_e3ad3cb40637df775f3a1548bdc12f21_1
lcon_380e33f297f7ed1f8979ee3c73775131_5
lcon_d54ae680c1f383c95f1a9409b4550c1d_1
lcon_6151c3028a05f391dd8dc379ddc1bd3a_1
lcon_046aacc1eca4b8d151c41766368c6e43_1
lcon_0d2225f00fe7d785f57724eba97ce882_2
lcon_b342c2893dc9ff4153307b651ba507eb_2
lcon_e8270051a4df57d5654efa9ac6c7dc25_1
lcon_5f57ebc7d7173f1fe78e816947995184_7
lcon_894ca811afb6e85b84b83ed24f354bb3_1
lcon_24131f67c4569d2487908e1d7c833624_13
lcon_b515e06dc3bce92bee8b9b6aff313cdb_3
lcon_28ac4701af388ef42f55f74aa700398d_1
lcon_9ca7622361a25e61a0adc285719ff18f_1
lcon_14eea7f551701c0e5c2553f221c024fd_**37**
lcon_0cad7738684b45ff74f27f27a09a441b_1
lcon_01645ac6a44649114330fbaf63679c50_8
lcon_7b34c75f6624268127c982060fe8cd97_1
lcon_9464efacf9d9d6e150291ab7dce3aee9_1
lcon_a6ec8080a035b83541acd43eb3b2eee0_1
lcon_f30e9b9879e0ae7231553a877694b7c1_1
lcon_fd916b3ba0ce897505bb73f7d95b8603_1

Clade I
Adriatic Sea and
Ionian Sea
Clade II
Clade III
Clade IV
Southern Ocean
Clade V
Clade VI
Clade VII
Clade VIII
Clade IX
Clade X
Clade XI

0.01 substitutions/site

Fig. 3.5. Neighbour joining tree illustrating the relationship among the V9-based secondary clusters from the *Tara* Oceans dataset along with *Leptocylindrus* species, *Tenuicylindrus belgicus,* other diatom species and few autotrophic stramenopiles. Bootstrap values have been generated with 1000 replicates. Bootstrap support for clades receiving <50% support have been omitted.

cluster with 38943 sequences, whereas other secondary clusters contained just 2-25 sequences.

Clade III was found to be composed of a single end-node consisting of a secondary cluster of 15 sequences. Although the distance criterion assigned this end node to *L. aporus*, it did not contain any reference sequence.

Clade IV included 13 end-nodes of which four were singletons, two secondary clusters contained pairs of sequences, and the remainder included five to 35 sequences. Although the distance criterion assigned all the sequences in Clade IV to *L. aporus*, it did not contain any reference sequence.

Clade V included a single end-node, composed of a secondary cluster of five sequences. Although the distance criterion assigned these sequences to *L. aporus*, it did not contain a reference sequence.

Clade VI (55% bootstrap support) contained the reference sequences of *L. danicus* and *L hargravesii*, as well as 26 end nodes, most of which represented only one or a few sequences. Yet, four of the end-nodes contained 409, 132, 810, and 3252 sequences.

Clade VII included 66 end nodes, most of which represented only one or a few sequences. Yet, six of the end nodes consisted of secondary clusters that contained 55 or more sequences (134, 245, 144, 201, 55, 283). Although the distance criterion assigned all the sequences in this clade to *L. aporus*, it did not contain any reference sequence.

Clade VIII included 11 end-nodes of which seven contained 6 or more sequences, the most populous of which contained 625 sequences. Although the distance criterion assigned all the sequences in this clade to *L. convexus*, it did not contain any reference sequence.

Table 3.6. Summary of Tara Oceans NJ trees. The number of end-nodes within each clade, their total sequence abundance and the geographical origin of the sequences are presented.

Clade	No of end nodes	Sequence abundance	Reference species	Geographical location
Clade I	15	288	-	East African coast, Mediterranean sea, Indian Ocean
Clade II	112	39110	*L. aporus*	multiple sampling stations
Clade III	5	15	-	multiple sampling stations
Clade IV	16	165	-	South Atlantic-, South Pacific- Ocean, Arabian-, Red- Sea, Antarctic Peninsula
Clade V	3	5	-	Southern Ocean, few African Coast and Arabian Sea
Clade VI	54	4666	*L. danicus/hargravesii*	multiple sampling stations
Clade VII	134	2090	-	multiple sampling stations
Clade VIII	51	824	-	multiple sampling stations
Clade IX	1	1	*L. minimus*	Alboran Sea
Clade X	15	79	-	multiple sampling stations
Clade XI	113	30367	*L. convexus*	multiple sampling stations

Clade IX included the reference sequence of *L. minimus* and was sister to an end-node of a singleton. However, this singleton sequence was assigned to *L. convexus* because its dissimilarity to the reference sequence of this species was smaller than that to the reference sequence of *L. minimus*. Nevertheless, from a phylogenetic viewpoint it grouped with *L. minimus*, though without any bootstrap support.

Clade X contained three end nodes, two of which were singletons and one contained 77 sequences. Although the distance criterion assigned all the sequences in this clade to *L. convexus*, it did not contain any reference sequence.

Clade XI contained the reference sequence of *L. convexus*, as well as 70 end nodes, 49 of which were singletons. The end-node closest to the reference sequence contained 30135 sequences. The distance criterion assigned all the sequences in this clade to *L. convexus*.

3.3.4 Biogeography. *Leptocylindrus* sequences were detected in 32 out of the 35 stations sampled and sequenced by the *Tara* Oceans team. More stations were sampled, but the sequence results were made available following the analysis carried out in this study. Distribution of clades sequences across the stations is presented in Fig. 3.6 and their abundance in different size fractions is presented in Table 3.6, Table 3.8 and Table 3.9. Geographical demarcation in clades was observed (see Fig. 3.5 and Table 3.6, Table 3.8 and Table 3.9). Within Clade I, Clade I-D was found exclusively in the Adriatic and Ionian Seas, and all sequences of Clade IV were found near the Antarctic Peninsula except for a few sequences from the station 66 (8 sequences) near Cape Town and station 36 (1 sequence) in the Arabian Sea. Within Clade VI, the *L. danicus* and *L. hargravesii* clade the Sub-clade VI-A (1357 sequences) appeared to be occurring in temperate latitudes and Sub-clade VI-B (3390 sequences) was found in non-temperate latitudes. Other clades contained sequences from multiple stations from all over the expedition track.

3.4 Discussion

Discrimination of taxa in recent methods such as metagenomics generally depends on the choice of the genetic marker and cut-off value used for taxon discrimination. A growing dataset of reference SSU rRNA sequences, covering virtually every known protist taxonomic group has led to its development as the marker of choice for estimation of diversity in sequence based surveys in environmental samples (Medlin and Kooistra 2010). Despite the apparent advantages of availability of diverse primers and hyper-variable regions that provide phylogenetic signals, there are limitations to its application. For example, the copy number of rRNA molecules varies between different evolutionary lineages and species, which make it challenging to evaluate whether a library has captured a large enough fraction of the diversity and to further estimate the relative abundance of different phylotypes in the source environment (Case *et al.* 2007). Another limitation lies in the discrimination level offered by the variable regions, as the regions mostly categorize thousands of closely related sequences (clades of secondary clusters) that have hardly been linked to already sequenced morphospecies (Caron 2009, Nebel, 2011). Risk in interpreting the species richness and diversity though SSU rDNA tags in a sample is also due to the fact that not all the SSU rDNA gene copies are necessarily identical (Rooney 2004). Additionally, at times even the hyper-variable regions have failed to discriminate the closely related species (Stoeck *et al.* 2010). The V4 and V9 regions of the nuclear encoded SSU rDNA genes, most commonly used SSU rDNA regions in the metagenomic study of diversity and to distinguish different phylotypes, offer different levels of phylogenetic signals. The BioMarKs V4-tree reveals a series of well-supported clades, each of which contain one or two reference sequences of *Leptocylindrus* and *Tenuicylindrus*. The V4 region is the larger in size (≈330bp), shows a higher number of parsimony informative site and higher signal to noise ratio than the V9 region (≈130bp). This is also illustrated in the tree topologies of the V4 and V9 regions. The one based on

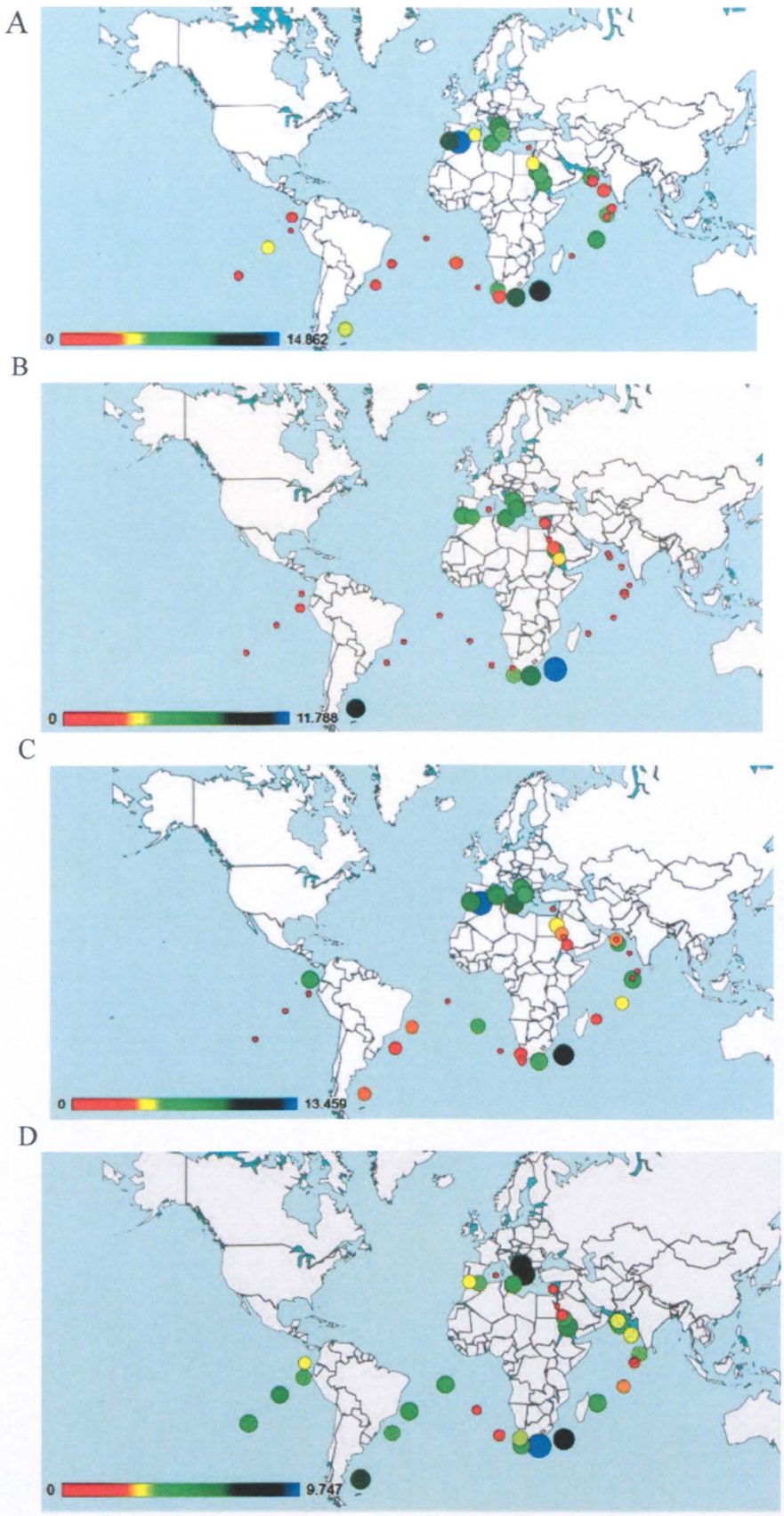


Fig. 3.6. Distribution of *Leptocylindrus spp.* and Clade VII inferred from the phylogenetic analysis of V9 metagenomic database of the *Tara* Oceans. (A) *Leptocylindrus aporus.* (B) *Leptocylindrus convexus.* (C) *Leptocylindrus danicus.* (D) Clade VII. Sequence abundance in $\log_{10}$ values.

the V4 was better resolved than the one based on the V9. These results are consistent with other studies of comparative morphological and marker analysis and estimates of species diversity and richness estimates. For example the comparative study of 29 species of the order Tintinnida concluded that V9 was less effective than V4 in taxon discrimination (Santoferrara *et al.* in press). In diatoms, many of the *Pseudo-nitzschia sensu lato* species cannot be distinguished at even at whole 18S rDNA and many of them differ at very few bases (Valeria Ruggiero personal communication). Hence in such cases both the regions fail to identify the species richness.

Singletons are considered to be the single biggest cause of over-estimation in species richness tests in environmental samples (Li *et al.* 2012). Many singletons and doubletons were observed in the datasets presented in this study. The swarms of singletons that are recovered within clades probably represent sequences that are a few point mutations away from others, whereas those that form markedly distinct clades on their own, e.g., the singleton that is sister to *L. minimus* in Clade IX in the *Tara* Oceans dataset and the singleton in Clade IV of the BioMarKs dataset, probably represent rare sequences of novel species. In theory, they might also be chimaeras due to a PCR error, but we screened these sequences and did not find evidence for them being chimaeras.

Singletons or errors in PCR sequencing also depend on the sequencing platform used. Each claims a high level of accuracy but still errors in sequencing are of major concern especially in environmental sample sequencing where it is difficult to understand whether the diversity in sequences is the diversity of the species in natural environment. For instance in a comparative study Illumina yielded longer and more accurate contigs despite the substantially shorter read length relative to Roche 454 and comparable average sequencing errors in the raw reads of the two platforms (0.5% per base). Nevertheless, the field of NGS is rapidly progressing and advanced and efficient methods are continuously

under replacements. Errors are also introduced during the PCR reaction therefore if amplicon sequencing (involving PCR reaction before DNA sequencing) is followed over direct DNA sequencing then there are more chances of errors being incorporated into the database thus generating a false diversity value for the sample. To eliminate errors algorithms like SLP (Huse *et al.* 2010), PyroNoise (Quince *et al.* 2009), Denoiser (Reeder and Knight 2010) and Ampliconnoise (Quince *et al.* 2011) were developed that focus at identifying and removing sequence noise. Thus, species richness is an arbitrary value, which may or may not reflect the actual diversity in the environment. Despite the problem in handling singletons in the dataset and in the phylogenetic trees, we observed considerable diversity in the studied species.

Clone libraries generated through isolation of single cells or through PCR amplification by species specific primers provide another window for diversity estimates in phytoplankton. The level of diversity interpreted from clone libraries likewise the metagenomic studies with NGS, depends on the choice of marker and the effort made in sequencing (sample size).

Dissimilarity cut-off values. The diversity inferred in the natural samples inferred through the metagenomic sequencing depends on the choice in dissimilarity cut-off value chosen for sequence retrieval and further taxonomic identity assignment. Higher sequences dissimilarity may result in the low number of species diversity and a lower cut-off may result in higher species diversity. This choice depends on the molecular marker used for the estimation of the diversity and on the distance of separation between the morphospecies.

3.4.1 ***Support for the different recognized species in Chapter II.*** The nuclear SSU rDNA was able to distinguish all the species in *Leptocylindrus* including *L. danicus* and *L. hargravesii.* However, the latter two species differed only in a few bases in the V4 region

and in none at all in the V9 region. It is therefore not surprising that within the phylogeny of the *Tara* Oceans-dataset the marker sequences of these two species are found together in a clade. However, what is surprising is that the reference sequences of these two species are embedded within a larger clade with many end-nodes, several of which are composed of secondary clusters, each with large numbers of sequences. In fact, the two reference sequences are resolved in one of these clusters. Within the BioMarks V4-dataset, a similar pattern is revealed. The reference sequences of the two species are not identical, but at least closely related. The reference sequences are embedded within a group of closely related sequences, some of which happen to be more similar to that of *L. hargravesii* and others to that of *L. danicus*. So, the taxonomic assignment of these query sequences based on dissimilarities to these two marker sequences is rather arbitrary. A possible interpretation for all this sequence diversity within *L. danicus – L hargravesii* is that the secondary clusters with their many sequences represent distinct populations or even distinct species within a *L. hargravesii / L. danicus* species complex. All this does not questions the validity of the two species as established in Chapter II because Neapolitan strains of these two species formed two groups based on differences on five molecular markers.

In the BioMarKs V4 dataset, the single secondary cluster of *L. convexus* suggests that this morphological species consists of a single species. This conclusion is of course valid only within the restricted geographical distribution in the GoN. Within the *Tara* Oceans V9, *L. convexus* is found in Clade XI with considerable microvariation, but no subdivision in distinct clades, and therefore, no clear evidence either for multiple species within this morphological taxon.

The phylogeny inferred from the BioMarKs V4 dataset suggests that *L. minimus* consists of a single species because the sequences in its clade show no clear further subdivision into well-supported clades. The absence of any query sequences very close to *L. minimus* in the

phylogeny inferred from the *Tara* Oceans V9 dataset, indicates that the species is absent in the regions sampled by *Tara* Oceans, that is, at least, during the sampling dates of the stations. However, the results do not prove the absence of the species in these regions because the species may occur at these stations in the plankton outside the sampling period. However, a singleton sequence was found as relative of the *L. minimus* reference sequence, but markedly dissimilar from it, suggesting that this sequence belongs to an individual that is either a variety of the species or constitutes a close sister species of *L. minimus.*

The phylogeny inferred from the BioMarKs V4 dataset suggests that *L. aporus* consists also of a single species because as in the previous species the sequences in its clade show no clear further subdivision into well-supported clades. Within the *Tara* Oceans V9, *L. aporus* is found in Clade II with considerable microvariation, but as in the previous species, no clear subdivision in distinct clades, and therefore, no clear evidence either for multiple species within this morphological taxon.

The phylogeny inferred from the BioMarKs V4 dataset suggests that *T. belgicus* consists also of a single species because also this clade shows no clear further subdivision into well-supported clades. Within the *Tara* Oceans V9, query sequences grouping with *T. belgicus* were unfortunately eliminated from the dataset.

3.4.2 ***Possible evidence for additional*** **Leptocylindrus** ***species.*** In the BioMarKs V4 dataset, Clade IV contains a singleton sequence, which might represent a species new to science. It must be a *Leptocylindrus* species because it is embedded within a clade containing other species belonging to this genus. The fact that it is just a singleton could be interpreted as it being a PCR error, specifically a chimaera. However, visual examination showed that base chances with the sequences of its closest sister sequences of *L. aporus* are distributed all over the sequence. This means that the sequence could be a pseudogene

(dysfunctional relatives). However, the differences are not scattered haphazardly all over the sequence but localized in small groups in regions where the differences among species are found. Therefore, this sequence is probably the representative of a rare species close to *L. aporus*.

Within the phylogeny inferred from the Tara Oceans V9 region Clades I, III, IV, V, VII, VIII and X do not contain any reference sequences. This implies that these clades belong to species new to science or they represent species for which no reference sequences have been collected yet. Alternatively, and given that the clades are not well supported, they might belong to *Leptocylindrus* species mentioned above or even to other centric genera. Yet, if it is assumed that the tree topology reflects true relationships among the *Leptocylindrus* sequences, then Clades III–XI together form a clade and all sequences inside this clade can be assumed to belong to *Leptocylindrus* because the marker sequences recovered within this clade all belong to *Leptocylindrus*. Clade I might then belong to another genus closely related to *Leptocylindrus* because it is close to radial centric outgroup taxa.

The lack of any additional clades close to these existing clades of these genera within the in the V4 tree suggests that there are no additional *Leptocylindrus* species in Western Europe. The alternative explanation is that, due to the cut-off value of 0.90, clades containing sequences that are still within *Leptocylindrus* but that represented different species were discarded in the selection procedure. The clades in the phylogeny inferred from the Tara Oceans V9 dataset are generally poorly supported or not at all. This is not surprising because the sequence is much shorter than the V4 region with only 104 parsimony informative positions.

The micro-variation in all of the clades in the V4 tree is comparable and so is the state of affairs in the V9 tree. However, the fact that there are no end-nodes on long branches inside the clades could be a consequence of the restrictive cut-off value of 0.90.

3.4.3 ***Biogeography.*** Results of the present study indicate that all *Leptocylindrus* species recognized in Chapter II are widely distributed. According to the *Tara* Oceans dataset some of the species appear to be more typical coastal whereas others are also well represented in mid oceanic samples. Yet, none of the species is exclusively coastal; they are present in low numbers also in mid-oceanic samples, suggesting that they occur there naturally or that they are being transported by ocean currents from one coastal region to the other on the opposite side of the ocean. Apparently, for *Leptocylindrus* species, everything can migrate everywhere, but the environment selects if the species can occur there in large numbers.

However, a few clades without marker sequences (Clades I and IV) reveal a more restricted distribution. For instance, Clade I is found only in the Adriatic and Ionian seas and not even in other regions within the Mediterranean Sea. So, apparently, a few clades are not everywhere. This observation fits observations in Kooistra et al. 2007 that a few *Skeletonema* species showed a restricted distribution pattern.

Despite the high number of sequences at many sites, absence of sequences in a sample from a particular geographical location does not necessarily mean that the species is absent. Some species show a restricted period of occurrence whereas some others exhibit longer periods of occurrence (Chapter II). Also phytoplankton occurrence is patchy, which challenges a conclusion that the species does not occur in a region because we did not detect it in a sample. Thus it remains difficult to quantify species richness through sequencing of SSU rDNA regions in metagenomic projects if the sampling is based on a single visit of a site, especially in regions showing strong seasonality.

Sampling sediments is often considered as alternatives to track the presence or absence of species and to construct their spatial distribution. Sediments are seeding store houses preserving material for the next bloom in the form of resting stages as spores, cysts, or

resting cells. Resting stages often result from sexual reproduction and are typically considered as traits to survive short-term and seasonal unfavourable conditions (Lewis 1999) although in most diatoms they are formed during the vegetative growth phase. Some studies have reported that resting stages can survive for years to decades in the natural environment. They are an important source of genetic novelty (Genovesi-Giunti *et al.* 2006) and also aid dispersal. It is often viewed that what is not observed in surface layers can be found in the sediments. The BioMarKs sediment samples contained mainly sequences from species that were also present in the surface layers, indicating that the fallen material was amplified and sequenced. In the Oslo sample of June 2010, *L. danicus* was detected in the sediment whereas in the September 2009 sample it was present in the plankton suggesting that this species over-summers in the benthos. This is true for all the sites including station Varna in Black Sea, where none of the studied species were found. However, phytoplankton checklists for Black Sea have readily provided long term records of *Leptocylindrus danicus* (Proshkina-Lavrenko 1955). Hence, failure of sequence retrieval in sediments does not necessarily imply that the species is absent at the site of sampling. Part of the reason for this might be that DNA extraction protocols result in low quantity of DNA from resting stages that are hard to break, anyway. The abundant DNA from freshly fallen diatom cell from recent blooms might be over amplified because of its higher quantity and easier accessibility. As a result, it is difficult but not impossible to construct the biogeographic distribution for a species. Certainly a spatial distribution by simply reporting as presence-absence based estimate of biodiversity can be readily assembled.

3.5 Conclusion

Understanding the diversity, biogeography and ecological role of protists depends on the degree of correlation between the morphological and molecular characters. Morphological analysis involves careful observations for cryptic species whereas estimating and

interpreting species richness from the metagenomic sequences of environmental samples depends on the choice made in the use of molecular marker and the cut-off value. Among the variety of choice available for genetic markers selection to estimate the diversity of species in environmental samples, SSU, V4 and V9 are the widely used, which provide similar taxon discrimination, while V9 was less effective in producing reliable distance trees. The diversity observed for *Leptocylindrus spp.* and *T. belgicus* was marginal with the V4 region with restricted geographical sampling in BioMarKs. On the other hand Tara Oceans with a greater sampling provides a contrasting view of the genetic diversity in the natural environment. The novel clades obtained can belong to any of the unknown genetic identity or even a novel species. However, given the uncertain phylogenetic trees generated with V9 region and without any information on the morphology of the identity, it is difficult to conclude the definitiveness of the novelty. Thus continued thoughts and efforts in the choice and improvements of the bioinformatic tools used for analysis of metagenomics data, can result in better interpretations of the species richness in environmental samples and finally to expand the studies of biogeography and ecology.

Table 3.7. Summary of V4 sequence abundance in BioMarKs sampling stations based on the assignment of reference species by CD-HIT-EST-2D clustering analysis.

Sample No.	Station Year	Depth	Size Fraction	Template	Total sequence	Clade I *L. danicus*	Clade I *L. hargravesii*	Clade II *L. convexus*	Clade III *L. minimus*	Clade V *L. aporus*	Clade VI *T. belgicus*
S0234	Barcelona 2010	Subsurface	0–0.2	DNA	6697					3	
S0155	Barcelona 2010	Subsurface	20–2000	cDNA	18349					1	
S0156	Barcelona 2010	Subsurface	20–2000	DNA	17934					2	
S0066	Barcelona 2010	Subsurface	3–20	cDNA	8902		1				
S0157	Gijon 2010	Subsurface	0.8–3	cDNA	15383					1	
S0158	Gijon 2010	Subsurface	3–20	cDNA	5685					5	
S0160	Gijon 2010	Subsurface	3–20	DNA	14130					43	
S0052	Naples 2009	DCM	0.8–3	cDNA	15866	1	6			1053	
S0161	Naples 2009	DCM	0.8–3	cDNA	19219	1	11			917	1
S0162	Naples 2009	DCM	0.8–3	cDNA	20905		5			250	
S0042	Naples 2009	DCM	0.8–3	DNA	10131		4			210	
S0167	Naples 2009	DCM	0.8–3	DNA	14121		4			252	
S0168	Naples 2009	DCM	0.8–3	DNA	15243		5			185	
S0235	Naples 2009	DCM	0–0.2	DNA	739		0			35	
S0041	Naples 2009	DCM	20–total	cDNA	8409	1	5			64	
S0164	Naples 2009	DCM	20–total	cDNA	23645		9			76	3
S0045	Naples 2009	DCM	20–total	DNA	9751		3			20	
S0170	Naples 2009	DCM	20–total	DNA	16022	1	7			21	
S0048	Naples 2009	DCM	3–20	cDNA	8637		25			3851	
S0165	Naples 2009	DCM	3–20	cDNA	29232	9	69			11782	5
S0049	Naples 2009	DCM	3–20	DNA	10053		21	1		2076	

Contd...

S0171	Naples 2009	DCM	3–20	DNA	15076	3	11			5399	
S0005	Naples 2009	Sediment	total	cDNA	20490	63	65			194	1
S0147	Naples 2009	Sediment	total	cDNA	21790	41	46			93	2
S0008	Naples 2009	Sediment	total	DNA	26540	13	7			29	
S0148	Naples 2009	Sediment	total	DNA	7985	3	1			8	
S0050	Naples 2009	Subsurface	0.8–3	cDNA	22249	4	2			1788	2
S0163	Naples 2009	Subsurface	0.8–3	cDNA	27024		7			1231	2
S0051	Naples 2009	Subsurface	0.8–3	DNA	8111	1	2			144	
S0169	Naples 2009	Subsurface	0.8–3	DNA	30255	1	3			456	3
S0236	Naples 2009	Subsurface	0–0.2	DNA	2343		3			18	
S0017	Naples 2009	Subsurface	20–total	cDNA	1715		1			21	1
S0125	Naples 2009	Subsurface	20–total	cDNA	15371	8	17			170	5
S0019	Naples 2009	Subsurface	20–total	DNA	1450					7	
S0135	Naples 2009	Subsurface	20–total	DNA	16454		14			3036	4
S0046	Naples 2009	Subsurface	3–20	cDNA	3646		5			1246	2
S0166	Naples 2009	Subsurface	3–20	cDNA	30868	8	42			8867	27
S0047	Naples 2009	Subsurface	3–20	DNA	14471	3	5			3043	1
S0172	Naples 2009	Subsurface	3–20	DNA	8064		4			1801	1
S0098	Naples 2010	DCM	0.8–3	DNA	8732					1	
S0086	Naples 2010	DCM	20–2000	cDNA	12968			1			
S0096	Naples 2010	DCM	20–2000	DNA	10936			1		1	
S0087	Naples 2010	DCM	3–20	cDNA	13001	2	3	1			
S0097	Naples 2010	DCM	3–20	DNA	8367	3	2	1		5	
S0089	Naples 2010	Sediment	total	cDNA	6972	5	1				
S0099	Naples 2010	Sediment	total	DNA	17061	5	3				

Contd...

S0085	Naples 2010	Subsurface	0.8–3	cDNA	8054			1		3	
S0095	Naples 2010	Subsurface	0.8–3	DNA	3812		1	3		20	
S0083	Naples 2010	Subsurface	20–2000	DNA	7338	1		1		2	
S0074	Naples 2010	Subsurface	3–20	cDNA	4413	4	2	3		6	
S0084	Naples 2010	Subsurface	3–20	DNA	3797	1	1	4		35	
S0025	Oslo 2009	DCM	0.8–3	cDNA	1978	2					
S0127	Oslo 2009	DCM	0.8–3	cDNA	11179	3	5	2	1	2	2
S0027	Oslo 2009	DCM	0.8–3	DNA	1180					1	
S0137	Oslo 2009	DCM	0.8–3	DNA	17176	20	10		1	14	
S0021	Oslo 2009	DCM	20–2000	cDNA	1098	8	1			3	
S0128	Oslo 2009	DCM	20–2000	cDNA	20585	52	20			15	4
S0138	Oslo 2009	DCM	20–2000	DNA	13533	8	3			6	
S0033	Oslo 2009	DCM	3–20	cDNA	1193	7	3		2	26	3
S0126	Oslo 2009	DCM	3–20	cDNA	9738	95	52	6	15	131	26
S0035	Oslo 2009	DCM	3–20	DNA	404	1				2	
S0136	Oslo 2009	DCM	3–20	DNA	13928	24	8	6	16	97	10
S0006	Oslo 2009	Sediment	total	cDNA	12458	2	7	2	1	1	1
S0010	Oslo 2009	Sediment	total	cDNA	8916	5	14				1
S0129	Oslo 2009	Sediment	total	cDNA	3971	3	2				
S0130	Oslo 2009	Sediment	total	cDNA	8128	8	8			4	1
S0007	Oslo 2009	Sediment	total	DNA	7195		3				
S0009	Oslo 2009	Sediment	total	DNA	3196		2				
S0139	Oslo 2009	Sediment	total	DNA	8300	1	1				
S0140	Oslo 2009	Sediment	total	DNA	15734	3	2			4	2
S0029	Oslo 2009	Subsurface	0.8–3	cDNA	2078	9				2	2

Contd...

S0131	Oslo 2009	Subsurface	0.8–3	cDNA	11063	40	11			7	6
S0031	Oslo 2009	Subsurface	0.8–3	DNA	424	1				2	
S0141	Oslo 2009	Subsurface	0.8–3	DNA	9494	32	10		5	2	1
S0065	Oslo 2009	Subsurface	20–2000	cDNA	10451	5	6			2	1
S0075	Oslo 2009	Subsurface	20–2000	DNA	9818	10	1		1	5	
S0037	Oslo 2009	Subsurface	3–20	cDNA	1506	6				2	2
S0132	Oslo 2009	Subsurface	3–20	cDNA	14666	86	37		9	31	37
S0039	Oslo 2009	Subsurface	3–20	DNA	1157	2	2		1	2	
S0142	Oslo 2009	Subsurface	3–20	DNA	16654	76	37		5	45	13
S0133	Oslo 2010	DCM	0.8–3	cDNA	7649			2	13		
S0143	Oslo 2010	DCM	0.8–3	DNA	13672			14	77		
S0105	Oslo 2010	DCM	3–20	cDNA	17313			8	18		
S0115	Oslo 2010	DCM	3–20	DNA	23504			2	43		
S0134	Oslo 2010	Sediment	total	cDNA	12535		1				
S0144	Oslo 2010	Sediment	total	DNA	15985	1	2				
S0106	Oslo 2010	Subsurface	0.8–3	cDNA	10313			4	2		
S0116	Oslo 2010	Subsurface	0.8–3	DNA	16128			4	6		
S0108	Oslo 2010	Subsurface	20–total	cDNA	11709				1		
S0117	Oslo 2010	Subsurface	3–20	DNA	17920			16	29		

DCM- Deep Chlorophyll Maximum

Table 3.8: Estimated relative abundance of each clade in the thirty six stations included in the study is shown as percentage of the total number of sequences generated for the respective station.

Station	Total	Clade I	Clade II	Clade III	Clade IV	Clade V	Clade VI	Clade VII	Clade VIII	Clade IX	Clade X	Clade XI
4	12032244		0.01203		0.00002		0.00044	0.00008				0.00304
7	13054106	0.00001	0.22807				0.00026	0.00011	0.00025	0.00001		0.08624
9	6504265		0.00046				0.00002	0.00061				0.00432
18	14545696		0.00253				0.00074	0.00132	0.00060		0.00010	0.00405
23	13243284	0.00141	0.00640				0.00046	0.00135	0.00002		0.00002	0.00152
25	12665073	0.00069	0.00047				0.00086	0.00002	0.00002		0.00029	0.00092
30	7469494		0.00001				0.00005	0.00004	0.00008			
31	4889342		0.00065					0.00039				0.00057
32	11360801		0.00443				0.00008	0.00055				0.00013
33	6631186		0.00112				0.00074	0.00072				
34	9353702		0.00346	0.00003		0.00001	0.00014	0.00012	0.00002			0.00006
36	13024346		0.00215					0.00018				0.00012
37	4266775			0.00007				0.00026				
38	10220216		0.00006		0.00001			0.00014				0.00066
41	14594147		0.00010					0.00003	0.00019		0.00005	
42	16728929	0.00002	0.00002			0.00001		0.00004	0.00001			0.00001
44	2210372		0.00294					0.00118				
45	9913410		0.00002				0.00003	0.00179				0.00169
48	3259809		0.01206				0.00003	0.02635	0.00006		0.00003	0.00058
52	15015336		0.00001					0.00016	0.00005			0.00006
64	10587727	0.00009	0.03086				0.03340	0.00011	0.00216		0.00012	0.04456
65	12350834		0.01123			0.00002	0.00198	0.00004	0.00056		0.00001	0.00098
66	10007905		0.00015		0.00008		0.00025	0.00002	0.00001			0.00003
67	6451556		0.00109					0.00045				0.00016
68	12947873		0.00001					0.00024	0.00022			
70	12090143		0.00013				0.00001	0.00013	0.00002			0.00042

72	22094164								0.00117		0.00000	
76	18180547		0.00002	0.00001				0.00065			0.00001	0.00008
78	19429802		0.00004					0.00001	0.00002			0.00005
82	17324392		0.00026	0.00001	0.00002		0.00237	0.00027				0.00008
84	8078955				0.00022							
85	22271842		0.00013		0.00059			0.00012				
98	15013819		0.00002				0.00001	0.00034	0.00001			
100	16995862		0.00018	0.00004			0.00001	0.00009	0.00023			
102	17597963						0.00001		0.00010			
109	23608165		0.00005					0.00004				0.00062

Table 3.9. Summary of V9 sequence abundance for each of the clades in the *Tara* Oceans sampling stations inferred through phylogenetic analysis.

Stn	Depth		Size	Clade I	Clade II	Clade III	Clade IV	Clade V	Clade VI	Clade VII	Clade VIII	Clade IX	Clade X	Clade XI
4	DCM	DNA	0.8-5		192				40	4				48
4	DCM	WGA/DNA	5-20		465					2				22
4	DCM	WGA/DNA	180-2000		7		2			2				1
4	SUR	DNA	5-20		639				4					270
4	SUR	DNA	0.8-5		143				9	2				25
4	SUR	DNA	180-2000		1									
7	DCM	DNA	0.8-5		1182				6					372
7	DCM	DNA	180-2000		6									3
7	DCM	DNA	20-180		987				7					266
7	DCM	WGA/DNA	5-20		4923				3	15				1953
7	SUR	DNA	0.8-5		888				12		1			167
7	SUR	DNA	180-2000		3							1		
7	SUR	DNA	20-180		1059				1		1			286
7	SUR	WGA/DNA	5-20	1	20725				5		30			8477
9	DCM	DNA	0.8-5		9				1					34
9	DCM	DNA	180-2000											
9	SUR	DNA	0.8-5		19									10
9	SUR	DNA	180-2000		2									5
18	DCM	DNA	5-20		3				5	1	2			7
18	DCM	DNA	0.8-5		25						1			3
18	DCM	DNA	180-2000						3	2	2			3
18	DCM	DNA	20-180		42				25	22	51		6	111
18	SUR	DNA	5-20		178				56	14	28		8	475
18	SUR	DNA	0.8-5		79				14	1				64
18	SUR	DNA	180-2000											
18	SUR	DNA	20-180		41				5		3			16
23	DCM	DNA	5-20	4	3						1			
23	DCM	DNA	0.8-5	74	562					5				54
23	DCM	DNA	180-2000	1	3				3	4				2
23	DCM	WGA/DNA	20-180		105				33	172				21
23	SUR	DNA	5-20	70	87				18	1				74
23	SUR	DNA	0.8-5	34	28									3
23	SUR	DNA	180-2000											
23	SUR	DNA	20-180	4	60				7	10	1		2	40
25	DCM	DNA	0.8-5	15	3									3
25	DCM	DNA	180-2000						1	2				
25	DCM	DNA	20-180	4	27				46	148	2		30	12
25	DCM	WGA/DNA	5-20	46	21				12	2			1	
25	SUR	DNA	5-20	10	4				35	3			6	16
25	SUR	DNA	0.8-5	12	5									1

Contd...

25	SUR	DNA	180-2000											
25	SUR	DNA	20-180						15	24	1		2	11
30	DCM	DNA	0.8-5		1				2					
30	DCM	WGA/DNA	20-180							1	2			
30	SUR	DNA	0.8-5											
30	SUR	DNA	180-2000							1				
30	SUR	DNA	20-180						2		4			
31	SUR	DNA	0.8-20		24									2
31	SUR	DNA	180-2000		2									
31	SUR	DNA	20-180		6									
32	DCM	DNA	0.8-5		31				3					5
32	DCM	DNA	180-2000											
32	DCM	DNA	20-180		44				5		2			3
32	DCM	WGA/DNA	5-20		414					1				
32	SUR	DNA	0.8-5		6									3
32	SUR	DNA	180-2000		5									
32	SUR	DNA	20-180		3				1	1				3
32	SUR	WGA/DNA	5-20							1				1
33	DCM	WGA/DNA	0.8-inf		41				45	18				67
33	SUR	DNA	0.8-20		33				4	1				4
33	SUR	DNA	180-2000											
33	SUR	DNA	20-180											
34	DCM	DNA	5-20		201				13	4				168
34	DCM	DNA	0.8-5		51					3				19
34	DCM	DNA	180-2000		1									
34	DCM	DNA	20-180		13			1		54	1			9
34	SUR	DNA	180-2000		1									
34	SUR	DNA	20-180		4						1			1
34	SUR	WGA/DNA	5-20		53	3				1				
36	DCM	DNA	0.8-5		42									388
36	DCM	DNA	180-2000		18					4				430
36	DCM	DNA	20-180		2					2				107
36	DCM	WGA/DNA	5-20		1					1				
36	SUR	DNA	0.8-5		99									1281
36	SUR	DNA	180-2000							7				14
36	SUR	DNA	20-180		104					21				2339
36	SUR	WGA/DNA	5-20		14					13				266
37	OMZ	WGA/DNA	0.8-5			3				1				
37	OMZ	WGA/DNA	20-180							10				
38	DCM	DNA	0.8-5							2				
38	DCM	DNA	180-2000											3
38	DCM	DNA	20-180											6
38	OMZ	WGA/DNA	5-20							16				4
38	SUR	DNA	0.8-5		2					1				9
38	SUR	DNA	20-180							1				5

Contd...

38	SUR	WGA/DNA	180-2000		4		1			4				
41	DCM	DNA	0.8-5							1				6
41	DCM	DNA	180-2000											
41	DCM	DNA	20-180							4	27		1	3
41	DCM	WGA/DNA	5-20		10					6				8
41	SUR	DNA	0.8-5											
41	SUR	DNA	180-2000											
41	SUR	DNA	20-180		4								6	51
41	SUR	WGA/DNA	5-20											
42	DCM	DNA	0.8-5	2										13
42	DCM	DNA	180-2000											1
42	DCM	DNA	20-180											
42	DCM	WGA/DNA	5-20	2	4					13				
42	SUR	DNA	0.8-5											3
42	SUR	DNA	180-2000											
42	SUR	DNA	20-180					1		1				4
42	SUR	WGA/DNA	5-20								1			
44	SUR	DNA	0.8-20		65					1				
45	SUR	DNA	5-20						3	3				
45	SUR	DNA	0.8-5		2					1				
45	SUR	DNA	180-2000											
45	SUR	DNA	20-180											
48	SUR	DNA	0.8-20		393				1	7	2		1	147
48	SUR	WGA/DNA	180-2000											
52	DCM	DNA	0.8-5											
52	DCM	DNA	180-2000											
52	DCM	DNA	20-180							1	8			
52	DCM	WGA/DNA	5-20							21				
52	SUR	DNA	0.8-5							1				1
52	SUR	DNA	180-2000											
52	SUR	DNA	20-180								2			
52	SUR	WGA/DNA	5-20		1					3				
64	DCM	DNA	0.8-5		176				216		2			356
64	DCM	DNA	180-2000		2					1				3
64	DCM	DNA	20-180	9	113				115	98	127		8	550
64	SUR	DNA	5-20		2680				2745	11				8263
64	SUR	DNA	0.8-5		255				425	2	1			569
64	SUR	DNA	180-2000		1						1			6
64	SUR	DNA	20-180		40				35	65	98		5	251
65	DCM	DNA	0.8-5		738				122	17	1			431
65	DCM	WGA/DNA	5-20		90				1	565				97
65	MES	DNA	0.8-180		2									
65	SUR	DNA	0.8-5		370				91	3				343
65	SUR	DNA	180-2000											
65	SUR	DNA	20-180		157			3	31	261	68		1	326

Contd...

65	SUR	WGA/DNA	5-20		30					12				459
66	DCM	DNA	0.8-5							1	1			
66	DCM	WGA/DNA	5-20		13				25	14				4
66	SUR	DNA	0.8-5							1				5
66	SUR	DNA	180-2000							1				
66	SUR	DNA	20-180							5				
66	SUR	WGA/DNA	5-20		2		8			2				2
67	SUR	DNA	5-20		7					1				4
67	SUR	DNA	0.8-5		13					1				46
67	SUR	DNA	0.8-inf		50					9				101
67	SUR	DNA	180-2000							1				
67	SUR	DNA	20-180											
68	DCM	DNA	0.8-inf							2	2			
68	DCM	DNA	180-2000											
68	DCM	WGA/DNA	5-20							1	12			
68	MES	DNA	0.8-3		1									
68	MES	DNA	3-inf											
68	SUR	DNA	0.8-5							1				1
68	SUR	DNA	20-180								14			
68	SUR	WGA/DNA	5-20							1				
70	MES	WGA/DNA	0.8-3		15				1	2				
70	MES	WGA/DNA	3-inf		1									
70	SUR	DNA	5-20											
70	SUR	DNA	0.8-5											
70	SUR	DNA	0.8-inf											
70	SUR	DNA	180-2000											
70	SUR	DNA	20-180								2			
72	DCM	DNA	5-20								7			
72	DCM	DNA	0.8-5											
72	DCM	DNA	0.8-inf							2	2			
72	DCM	DNA	180-2000											
72	DCM	DNA	20-180							1	20			
72	MES	WGA/DNA	0.8-3							17				16
72	MES	WGA/DNA	3-inf											
72	SUR	DNA	0.8-5							2				
72	SUR	DNA	0.8-inf							1	30			
72	SUR	DNA	180-2000							1				
72	SUR	DNA	20-180							1	200		1	
72	SUR	WGA/DNA	5-20							4				5
76	DCM	DNA	5-20											
76	DCM	DNA	0.8-5			1				31				54
76	DCM	DNA	0.8-inf											2
76	DCM	DNA	180-2000											
76	DCM	DNA	20-180											
76	MES	WGA/DNA	0.8-3											

Contd...

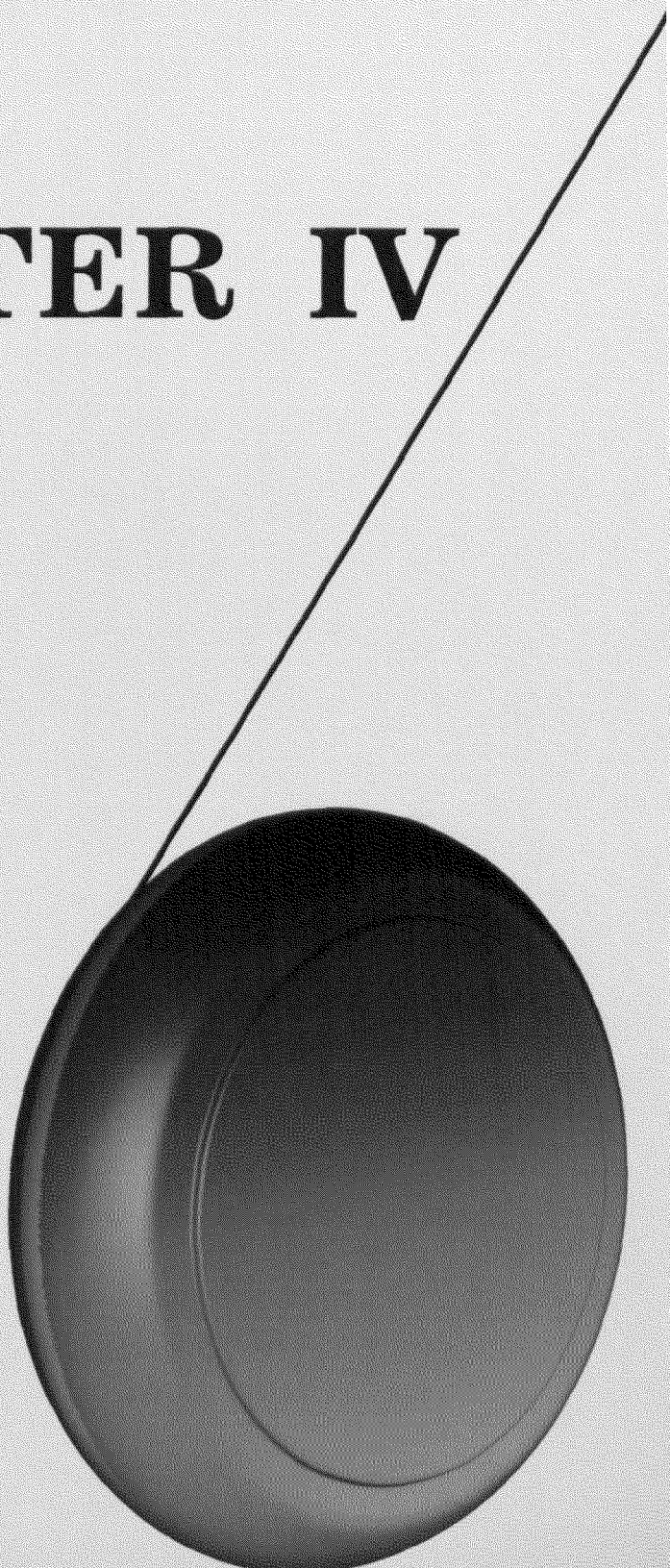
76	MES	WGA/DNA	3-inf		2									
76	SUR	DNA	5-20		1								1	2
76	SUR	DNA	0.8-5		1									35
76	SUR	DNA	0.8-inf											
76	SUR	DNA	180-2000											
76	SUR	DNA	20-180											
78	DCM	DNA	0.8-5		1									9
78	DCM	DNA	0.8-inf		1									7
78	DCM	DNA	180-2000											
78	DCM	DNA	20-180		2					2	1			1
78	MES	DNA	0.8-3							4				
78	MES	DNA	3-inf							1				
78	SUR	DNA	5-20							2	1			8
78	SUR	DNA	0.8-5							2				1
78	SUR	DNA	0.8-inf		2					2				10
78	SUR	DNA	180-2000											
78	SUR	DNA	20-180		1					3	1			3
82	DCM	DNA	0.8-inf				1		10	1				
82	DCM	DNA	180-2000											
82	DCM	WGA/DNA	5-20							17				
82	DCM	WGA/DNA	0.8-5		34	1				49				60
82	DCM	WGA/DNA	20-180		10					46				45
82	SUR	DNA	5-20						57					
82	SUR	DNA	0.8-5				2		153	2				3
82	SUR	DNA	0.8-inf		1				188					
82	SUR	DNA	180-2000				1		2	2				
82	SUR	DNA	20-180							1				
84	SUR	DNA	0.8-5				1							
84	SUR	DNA	0.8-inf				17							
84	SUR	DNA	180-2000											
84	SUR	DNA	20-180											
84	SUR	WGA/DNA	5-20							1				
85	DCM	DNA	0.8-5		2		25			1				
85	DCM	DNA	0.8-inf				6							
85	DCM	DNA	180-2000											
85	DCM	WGA/DNA	5-20				4			15				10
85	DCM	WGA/DNA	20-180		4		5			8				14
85	MES	WGA/DNA	0.8-3							2				
85	MES	WGA/DNA	3-inf		22					11				
85	SUR	DNA	0.8-5				50							
85	SUR	DNA	0.8-inf				32							
85	SUR	DNA	180-2000											
85	SUR	WGA/DNA	5-20				2			9				
85	SUR	WGA/DNA	20-180				8							
98	DCM	DNA	180-2000		1									

Contd...

98	DCM	DNA	20-180							1	2			
98	DCM	WGA/DNA	0.8-5							23				
98	MES	DNA	0.8-3							1				
98	MES	DNA	3-inf						1					
98	SUR	DNA	5-20		2					1				1
98	SUR	DNA	0.8-5											
98	SUR	DNA	180-2000							1				
98	SUR	DNA	20-180											
100	DCM	DNA	0.8-inf							2	23			
100	DCM	DNA	180-2000											
100	DCM	DNA	20-180								16			
100	DCM	WGA/DNA	5-20		30	7				42				19
100	MES	DNA	0.8-3						1	1				
100	MES	DNA	3-inf		1									
100	SUR	DNA	5-20							1				
100	SUR	DNA	0.8-5							2				
100	SUR	DNA	0.8-inf							1				
100	SUR	DNA	180-2000											
100	SUR	DNA	20-180							2				
102	DCM	DNA	5-20							5				
102	DCM	DNA	0.8-5							3				2
102	DCM	DNA	0.8-inf											
102	DCM	DNA	20-180						2	4	3			7
102	MES	DNA	0.8-3							1				
102	MES	DNA	3-inf							1				
102	SUR	DNA	5-20								1			1
102	SUR	DNA	0.8-5							1	1			
102	SUR	DNA	0.8-inf								10			
102	SUR	DNA	180-2000											
102	SUR	DNA	20-180								3			
109	DCM	DNA	0.8-5		2									3
109	DCM	DNA	0.8-inf							3				
109	DCM	DNA	180-2000							1				
109	DCM	DNA	20-180											1
109	DCM	WGA/DNA	5-20											
109	MES	DNA	0.8-3											
109	MES	DNA	3-inf							1				
109	SUR	DNA	5-20							5				1
109	SUR	DNA	0.8-5											
109	SUR	DNA	0.8-inf		8									1
109	SUR	DNA	180-2000		1									
109	SUR	DNA	20-180											

SUR-Surface, DCM-Deep Chlorophyll Maximum, WGA-Whole genome Amplication, MES- Mesopelagic
OMZ-Oxygen Minimum Zone

CHAPTER IV



Diversity of oxylipins and their biosynthetic pathways in the centric genera, *Leptocylindrus* and *Tenuicylindrus*

4.1. Introduction

Diatoms are a major component of the marine coastal ecosystems, forming the basis of the marine food chain and carrying out approximately one fifth of the photosynthesis on the Earth. These highly diverse photosynthetic organisms are found worldwide and populate almost all the aquatic ecosystem where sufficient light and nutrients are present to support their growth. It is estimated that around 200,000 species exists, of which approximately around 10,000 being currently described (Alverson 2008). The majority of non-described species are uncultivable and many are difficult to distinguish morphologically that are referred to as cryptic or pseudo-cryptic species.

Cryptic diversity has been well documented in the pennate diatoms (for example the genus *Pseudo-nitzschia* (Lundholm *et al.* 2002, Lundholm *et al.* 2003, Lundholm *et al.* 2006, Amato *et al.* 2007)), but the recent discoveries in centric diatoms (for example, *Skeletonema costatum* sensu lato (Sarno 2005, Sarno 2007; *Leptocylindrus,* Chapter II) have raised doubts on the factual diversity of diatoms. In recent years, a combination of different methods, involving microscopy and molecular markers (DNA sequences), have been applied to distinguish and identify diatom species. Metabolomics for chemical fingerprinting of species and functional genomics have proven particularly useful in plants and animals; however, its application in the diatom systematics is relatively less explored.

Evolved as the product of natural selection, secondary metabolites are the low molecular weight organic compounds that occur within the cell or tissue and represent a large group of structurally and functionally diverse chemicals. Microalgae produce a great variety of secondary metabolites that are synthesized by a diverse set of enzymes,

from a limited number of precursors from primary metabolism and possess a wide variety of biological function.

A major driving force in structuring the chemical diversity of secondary metabolites is the interactions between the organism and its environment. Although the functional analysis of secondary metabolites remains largely uninvestigated, a growing body of evidence suggests that chemical signals are involved in many processes, such as defense against grazers, competition, resource foraging and reproduction (Hay 2002). For example, compounds produced by microalgae such as dimethylsulfide (DMS) and its precursor dimethylsulfoniopropionate (DMSP) have important physiological functions such as osmolyte (Dickson and Kirst 1987), antioxidant (Karsten *et al.* 1992) or cryoprotectant (Sunda *et al.* 2002). In addition, DMSP play important roles in biological interactions; DMS released upon cell lysis triggers a tail-flapping response in the copepod *Temora longicornis* that probably helps the animals in finding food (Steinke *et al.* 2006). Thus the success of a species to perpetuate in a dynamic ecosystem depends, among other factors, on the complexity and diversity of molecules it can produce.

Polyunsaturated aldehydes (PUA) and polyunsaturated fatty acids (PUFA), along with their derivatives, constitute an important class of compounds with diverse roles, produced in diverse organisms, ranging from bacteria, fungi, algae to plants and animals. This group of compounds is generally named 'oxylipins', which include "any cyclic or acyclic product derived by the incorporation of oxygen into the carbon chains of PUFAs" (Gerwick *et al.* 1991, Cutignano *et al.* 2011). Diatoms are also known to produce oxylipins, which appear to play complex roles in predator defence processes as well as in other inter- and intra-specific interactions. Specifically, in diatoms oxylipins include short-chain unsaturated aldehydes and hydroxyl-, keto- and epoxyhydroxy fatty acid derivatives (Fontana *et al.* 2007). The first biosynthetic study in diatoms by Pohnert (2000), demonstrated the synthesis of 9-oxonona-5Z,7E-dienoic acid from C_{20}

fatty acids in benthic diatom, *Gomphonema parvulum*. Later, in *Skeletonema costatum* the formation of octadienal and octatrienal by lipoxygenase (LOX) mediated oxidation of C_{16} fatty acids was demonstrated.

Diatoms produce either one PUA or a mixture of several PUAs in wound-activated processes (Wichard *et al.* 2005, Fontana *et al.* 2007), e.g. upon cell breaking due to either grazing or mechanical damage, that trigger a cascade of reactions leading to the production of hydroperoxy-, hydroxy-, keto-, oxo-, epoxy-alcohols and aldehydes (Wichard *et al.* 2005). Oxylipin biosynthesis involves, as the first committed step, the oxidation of fatty acids to hydroperoxides by LOXs (classified based on their positional specificity in the molecular oxygen introduction), a non-heme iron dioxygenase that adds molecular oxygen to the carbon chain of fatty acids. In plants, oxygenation typically involves a C_{18} chain of linoleic and linolenic acids (the most abundant PUFAs in plants; Andreou, 2009), while in diatoms it generally involves C_{16} and C_{20} PUFAs (Cutignano *et al.* 2011). Alternatively, hydroperoxides may be formed through chemical oxidation of fatty acids (Schneider *et al.* 2007). Oxidation of fatty acids at different positions in the carbon chain leads to the formation of the corresponding hydroperoxides, for example oxygenation at C-9 by 9-LOX leads to the formation of 9-hydroperoxy-derivatives of the substrate, that are further metabolised to diverse compounds. In diatoms, the most commonly synthesised oxylipins includes C_{10} 2E,4E /Z-decadienal (2E,4E/Z-DD) and 2E,4E/Z,7Z-decatrienal (Miralto *et al.* 1999), but also C_8 2E, 4E/Z-octadienal, 2E, 4E/Z-octatrienal and $C_7$2E, 4E/Z-heptadienal (d'Ippolito *et al.* 2002, Wichard *et al.* 2005).

The ecological and biological role played by oxylipins is relatively well understood in plants where they play diverse biological functions as second messengers and as anti-microbial, anti-insecticidal as well as antifungal compounds (Knight *et al.* 2001, Eckardt 2008). In diatoms, they are apparently involved in defence mechanisms against

grazing activity, as both infochemical and allelochemicals (Leflaive and Ten-Hage 2009).

Diatoms were considered to be valuable food for copepods and other grazing organisms, until the mid-1990s, when one of the first experimental observation by (Ianora and Poulet 1993) revealed that a diatom diet can decrease egg hatching rate in the copepod *Temora stylifera*. This observation was further supported by many laboratory studies where copepods were either fed with PUA-producing diatom cultures (Ianora *et al.* 2004) or were exposed to pure compounds (Caldwell *et al.* 2004). Field studies have also demonstrated reduced egg hatching rates, like in the case of *Acartia clausi,* which showed a 12% reduced egg hatching rate (Miralto *et al.* 1999) and of *Pseudo-calanus newmani* and *Calanus pacificus*, which showed reduced reproductive success during blooms of *Thalassiosira* spp. (Halsband-Lenk *et al.* 2005, Pierson *et al.* 2005). The bioactive compounds responsible for such effects were firstly characterised by Miralto *et al.* (1999) as 2E,4E-DD,2E,4E,7Z-decatrienal and 2E,4Z,7Z-decatrienal. Through laboratory and field studies, the anti-proliferative effect of diatom PUAs is now clear, but it is not restricted to copepods. For example an earlier study by Caldwell et al. (2004) showed reduced sperm motility in sea urchins under the effect of 2E,4E/Z-DD. In polychaetes, crude and pure extracts of 2E,4E/Z-DD from *Skeletonema costatum* was found to inhibit fertilization, embryogenesis and hatching success (Caldwell *et al.* 2002). The mode of action of PUAs has been fairly predicted to be due to the apoptotic activity during the developmental stages (Romano *et al.* 2010). Independent of PUA production, the enzymatic oxidation of *Chaetoceros affinis* and *C. socialis* PUFAs and consequent teratogenic effects (malformation of embryos) was recently demonstrated in copepods (*Nauplii*) (Fontana *et al.* 2007). The synthesis of fatty acid hydroperoxides (FAHs) and other oxylipins (e.g., PUAs) in specific steps of defence signalling, and through the indirect boost of the oxidative stress with synthesis of lethal radical

chemicals (highly reactive oxygen species) was shown to reduce copepod reproductive capacity (Fontana *et al.* 2007). Despite these studies and demonstrations, the factors affecting copepod reproduction and the roles played by PUAs and PUFAs derivatives in different biotic interactions remain largely unclear.

The role of PUAs is not restricted to the defence against grazers, as new functions have recently been discovered which have widened the list of biological activities in which they are involved. Casotti *et al.* (2005) first proposed a role of PUA as chemical signals in interspecies communication under unfavourable conditions. The study demonstrated that 2E,4E/Z-DD could induce an apoptotic-like degenerative process causing cell death in *Thalassiosira weissflogii*, a species not producing 2E,4E/Z-DD. The response to 2E,4E/Z-DD was also studied in *Phaeodactylum tricornutum*, where it includes a rapid accumulation of nitric oxide (NO) in subcellular compartments and an altered expression of superoxide dismutase and metacaspases, which are involved in stress response pathways (Wolfe-Simon *et al.* 2006). Ribalet *et al.* (2007) showed that the sensitivity to PUAs is species-specific, demonstrating a role of these compounds in the control of bloom development and termination. Therefore the knowledge of bioactive lipid derivatives, and of the diversity in their synthesis pathways across different lineages, is crucial to the understanding of the biotic interaction in the natural environments.

Oxylipins synthetic pathways and the compounds eventually synthesized are highly diverse in the photosynthetic eukaryotes. Although the level of functional diversity is scarcely studied, some attempts have been made to use these compounds in chemotaxonomy as a complementary or alternative tool to classify and identify cryptic species. Surprisingly, a high species-specificity has been observed in both centric and pennate diatoms (Wichard *et al.* 2005, Fontana *et al.* 2007). Thus oxylipins might be used to characterise and classify species, to identify the physiological differences and at

times as chemical markers for identification in case of cryptic and pseudo-cryptic species.

In the case of *Leptocylindrus*, where pseudo-cryptic/cryptic species were identified in this thesis research (Chapter II), a study was planned to develop a metabolomics-based approach to the taxonomy of the genus by applying targeted Liquid chromatography–mass spectrometry (LC-MS) metabolite profiling of oxylipins. This study was done in scientific collaboration with Dr Angelo Fontana and colleagues at *CNR-Instituto di Chimica Biomolecolare*, Naples, Italy. As samples, cultures of selected strains isolated for the species diversity study of *Leptocylindrus* and *T. belgicus* were used. The discrimination among species was hence based on differences in their metabolite profiles in strains isolated in a single geographical location. During the study, firstly the diversity of compounds produced was investigated to determine whether specific metabolites could be adopted as biomarkers in the chemotaxonomy of the genus *Leptocylindrus*. Secondly, a comparative analysis of the phyco-oxylipins synthetic pathways of each species was performed.

4.2. Methods

4.2.1. Cultures. Two strains of each of 5 species (*Leptocylindrus aporus, L. danicus, L. convexus, L. hargravesii* and *Tenuicylindrus belgicus*) were selected (Table 4.1) from the existing culture collection established for the study presented in Chapter II. Stock cultures were maintained in an incubator at 20 °C and on a 12:12h L:D (Light:Dark) photoperiod with an irradiance of about 40–60 µmol photons $m^{-2}sec^{-1}$ by periodic transfers into 50 ml polystyrene flasks with 30 ml of K culture medium (Keller et al. 1987). Before the start of the experiment, the growth cycle of the individual species was investigated to determine the time needed to reach the stationary phase. To this end, 200 ml flat polystyrene bottles containing 100 ml of K medium were inoculated with 50

$x10^3$ cells from a culture in the exponential phase and were allowed to grow under a 12:12h L:D photoperiod with an irradiance of about 100 µmol photons $m^{-2}sec^{-1}$at 20°C until the decline phase. For the determination of the growth cycle, one strain of each species was used.

To obtain the material for the oxylipin analysis, cells were inoculated from exponentially growing cultures into two 2 l spherical, flat bottomed, glass flasks with 1 l of K media at a final cell density of 500 cells ml^{-1}. The cultures were allowed to grow under 100 µmol photons $m^{-2}sec^{-1}$ of light with an L:D cycle, 12:12. Cell densities were estimated every alternate day by counting cell number till the end of stationary phase. At the beginning of the stationary phase (or at the end of the exponential phase) the cultures were centrifuged and the culture pellets were collected. Pellets from 300 ml culture in triplicates of each species strain were frozen until used.

Table 4.1. Strains used in the analysis.

Strain Information	Collection Date
Leptocylindrus aporus	
SZN-B727	03/08/2010
SZN-B764	18/11/2010
Leptocylindrus convexus	
SZN-B783	25/01/2011
SZN-B778	25/01/2011
Leptocylindrus danicus	
SZN-B707	15/02/2010
SZN-B714	21/04/2010
Leptocylindrus hargravesii	
SZN-B781	25/01/2011
SZN-B772	21/12/2010
Tenuicylindrus belgicus	
SZN-B739	02/10/2010
SZN-B755	19/10/2010

4.2.2. Oxylipins analysis. Culture pellets collected for oxylipin analysis were extracted and analysed according to the methods described in (d'Ippolito *et al.* 2009). Briefly,

cells were suspended in 1 ml of filtered sea water 0.22 μm, sonicated for 1 min and incubated at room temperature. After 20 min, 1 ml of acetone and 5 μg 16-hydroxy-hexadecanoic acid (16-OH, internal standard) were added. The sample was sonicated for 1 min and centrifuged at 2000 *g* for 5 min at 5°C. The pellet was resuspended in 2 ml of H_2O/acetone (1:1, v/v), sonicated 1 min and centrifuged at 2000 *g* for 5 min. The supernatants were combined and extracted twice with equal volume of CH_2Cl_2. The organic layers were combined, dried over dry Na_2SO_4, hence filtrated and evaporated at reduced pressure. The dry extract was methylated with ethereal diazomethane for 30 min, and evaporated under nitrogen flow.

Methylated extracts were dissolved in methanol to a final concentration of 1μg μl^{-1} and directly analysed by LC-MS. The mass spectrometry (MS) method was based on a micro-Quadrupole time-of-flight (micro-QToF) instrument equipped with an electrospray ionization (ESI) source in positive ion mode and a UV photodiode array (DAD) detector (scan range 205–400 nm) for a dual monitoring of the chromatographic runs. For ESI-QToF-MS/MS experiments, argon was used as collision gas at a pressure of 22 mbar. Chromatographic analysis was carried out on a reverse phase column (Phenomenex, C-18 Kromasil 4.6 x 250 mm, 100 Å) using a linear MeOH/ H_2O gradient 75/25 to 100/0 in 30 min with a column flow of 1 ml min^{-1}. One tenth of the column flow was channelled by a post-column split to the ESI^+ (Q-Tof) MS analyser and the remaining 9/10 to the UV DAD detector (Cutignano *et al.* 2011).

4.3. Results

4.3.1. ***Growth.*** Many of the secondary metabolites are expressed during the stationary phase of the growth cycle and hence it is important to determine the time taken to reach that phase with a given inoculum. Before the start of the experiment, the growth curve for each species was determined. A typical growth cycle for each species is shown in Fig. 4.1, based on cell counts. All the species grew exponentially for 6–8 days and then

entered the decline phase. The only exception was the species *T. belgicus,* which showed a fluctuating cell density for an extended period of time, maintaining a plateau after reaching a maximum cell density (8.30 x10^5cells ml^{-1}). *Leptocylindrus danicus* (1.98x10^5 cells ml^{-1}) and *L. hargravesii* (2.10 x10^5cells ml^{-1}) reached a similar density, which was lower as compared to other species. *Leptocylindrus aporus* (4.50x10^5 cells ml^{-1}) and *L. convexus* (5.10x10^5 cells ml^{-1}) showed similar maximum cell density at the end of experiment.

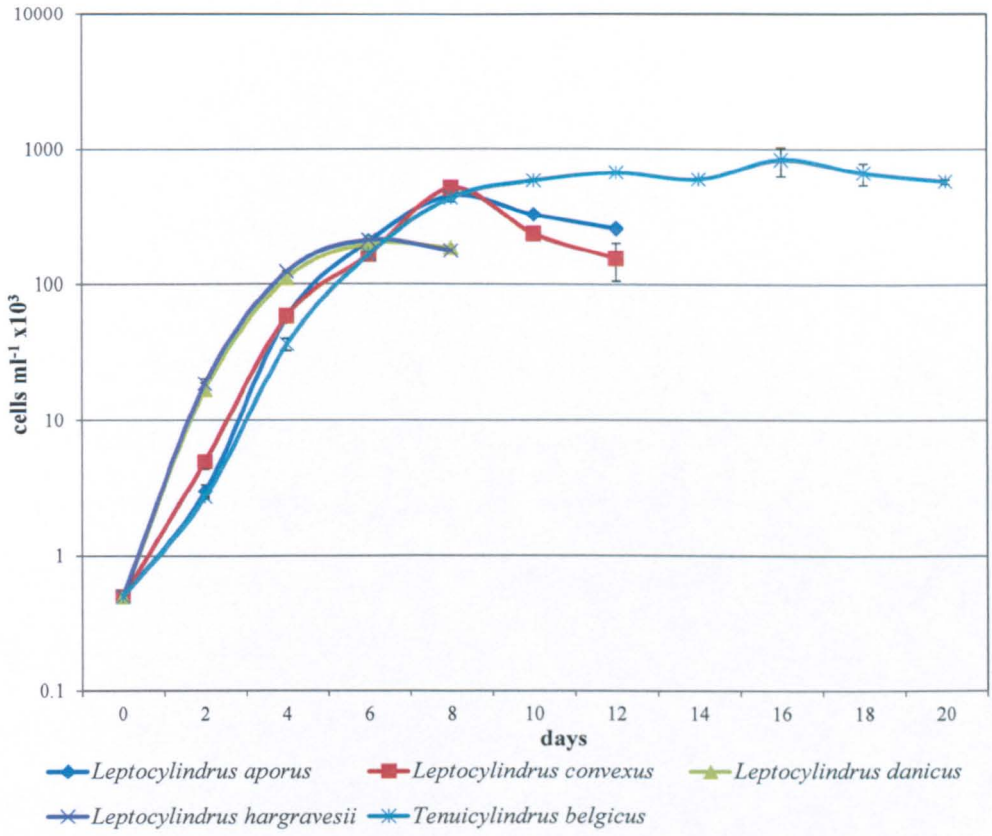


Fig. 4.1. Typical growth curves established by cell density for *Leptocylindrus* species *and T. belgicus* (n=3).

4.3.2. *Oxylipin profiles.* Using the standardised sample preparation protocol followed by the optimised LC-MS analysis, a reproducible chemical fingerprint of each species was obtained (Fig. 4.2 and 4.3). The mass (m/z), retention time (R_t), and UV absorption (λ_{max}) of major peaks in different samples are reported in Table 4.2. Strains of all the species showed similar oxylipin profiles with slight variation in the relative quantity.

Oxylipins profiles of *L. aporus* showed two major peaks on the chromatogram (Fig. 4.2; Table 4.2, A) with m/z 355.1 (R_t 19.3, λ_{max} 234) and 371.1 (R_t 14.5, λ_{max} 210). The later peak was abundant in the strain SZN-B727 in comparison with strain SZN-B764. *Leptocylindrus convexus* chromatograms (Fig. 4.2; Table 4.2, B) were characterised by the presence of peaks with m/z 381.2 (R_t 22.28, λ_{max} 234), 355.1 (R_t 19.32, λ_{max} 234), 397.2 (R_t of 22.28, λ_{max} 210), 395.2 (R_t of 22.28, λ_{max} 234), 371.1 (R_t 14.2, λ_{max} 210). The two strains, SZN-B783 and SZN-B778 had a similar profile except for the presence in SZN-B778 of a peak with m/z 353 (R_t 18.9, λ_{max} 234). *Leptocylindrus danicus* chromatograms (Fig. 4.2; Table 4.2, C) showed four major peaks with m/z 381.2 (R_t 22.48, λ_{max} 234), 355.1 (R_t 19.61, λ_{max} 234), 397.2 (R_t 17.9, λ_{max} 210), 371.1 (R_t 14.8, λ_{max} 210). In both strains, small quantities of isomers with m/z 371 were detected. The profile of *L. hargravesii* was characterised by presence of four major peaks (Fig. 4.2; Table 4.2, D and E) with m/z 381.2 (R_t 22.5, λ_{max} 234), 355.1 (R_t 19.6, λ_{max} 234), 397.1 (R_t 17.9, λ_{max} 210), 371.1(R_t 14.8 λ_{max} 210). *Tenuicylindrus belgicus* chromatograms (Fig. 4.3; Table 4.2, F) were constituted by two major peaks with m/z 355.1(R_t 21.3 λ_{max} 234) and 395.1 (R_t 19.0, λ_{max} 234), and additional minor peaks with m/z 371 at different retention time.

On the basis of the m/z, R_t, & λ_{max} and in comparison with the R_t of 16-OH (internal standard), few of the molecules were characterized. The compounds with m/z 355, λ_{max} 234, and R_t from 19 to 22 were assigned as hydroxyacids derivatives of eicosapentaenoic acid (EPA), while the molecules with m/z 371, λ_{max} 210, and R_t from 13 to 15 were characterized as epoxyalcohol derivatives of EPA. The compounds with m/z 381, λ_{max} 234, and R_t from 22 to 23 were assigned as hydroxyacids derivatives of docosa hexaenoic acid (DHA), while the molecules with m/z 397, λ_{max} 210, and R_t from 16 to 18 were characterized as epoxyalcohol of DHA.

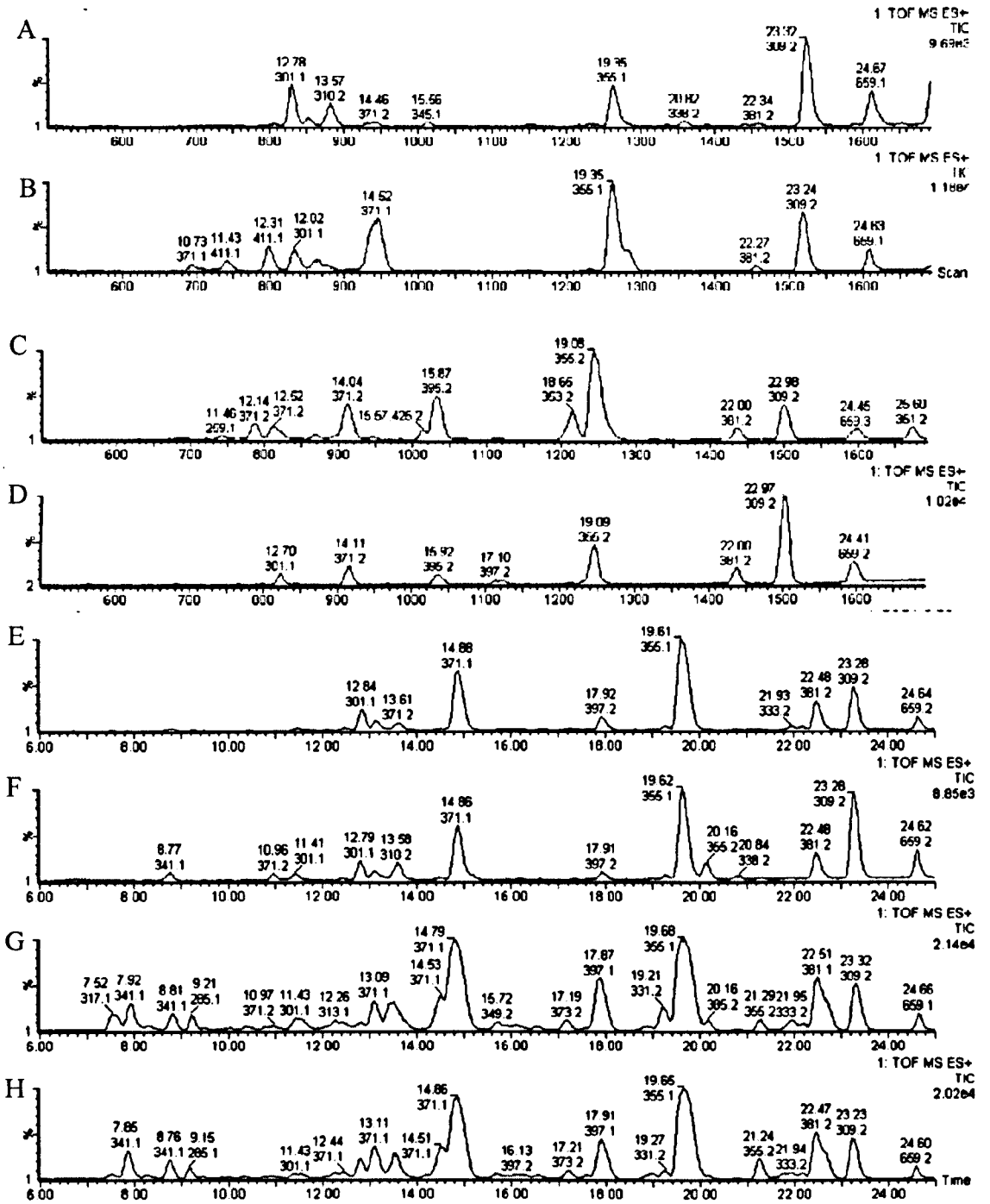


Fig. 4.2. Phyco-oxylipins signatures by LC-MS in *Leptocylindrus* species. (A) *L. aporus*, strain SZN-B764. (B) *L. aporus*, strain SZN-B727. (C) *L. convexus*, strain SZN-B778. (D) *L. convexus*, strain SZN-B783. (E) *L. danicus*, strain SZN-B707. (F) *L. danicus*, strain SZN-B714. (G) *L. hargravesii*, strain SZN-B772. (H) *L. hargravesii*, strain SZN-B781. Numbers above the peaks indicate the molecular ion ($M+Na^+$) as determined by ESI^+ ionization and retention time. 16-hydroxyhexadecaenoic acid methyl ester, shown by the mass peak at m/z 309.

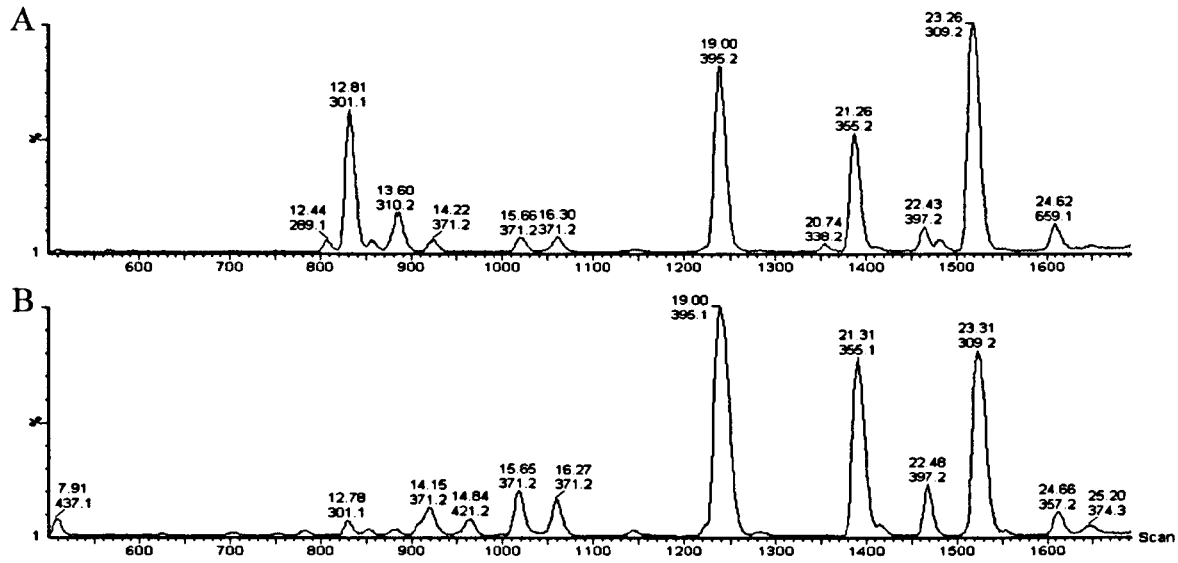


Fig. 4.3. Phyco-oxylipins signatures by LC-MS in *Tenuicylindrus belgicus*. (A) strain SZN-B739. (B) strain SZN-B755. Numbers above the peaks indicate the molecular ion ($M+Na^+$) as determined by ESI^+ ionization and retention time. 16-hydroxyhexadecaenoic acid methyl ester, shown by the mass peak at m/z 309.

4.3.3 ***Oxylipin characterisation.*** To characterize LOX pathways in diatom species, a simple and reproducible procedure recently published for the detection and characterization of oxylipins in microalgae (Cutignano *et al.* 2011) working directly on the extract was applied. The method was based on the fragmentation pattern of epoxyalcohol, generated at high collision energy using positive ESI^+ as source in a Q-ToF MS. If the hydroxyl moiety is between the epoxide and the carboxylic end, MS/MS fragmentation induces formation of ion clusters producing three main peaks that result by breaking of the carbon bonds of the epoxide, as well as the bonds between the epoxide and hydroxyl group. Contrastingly, if the hydroxyl moiety is between the epoxide and the methyl end, one detectable fragment is given that is derived from the cleavage of the epoxide ring. The mechanism is general and highly predictable. The analysis of the MS/MS spectra of epoxy-alcohols allows unambiguous prediction of the position of peroxidation, thus establishing the LOX activity responsible for the primary insertion of oxygen.

Table 4.2. Summary of the LC-MS chromatograms of *Leptocylindrus* species and *T. belgicus*. (A) *L. aporus*, strain SZN-B727 and SZN-B764 (B) *L. convexus*, strain SZN-B783 and SZN-B764. (C) *L. danicus*, strain SZN-B714 and SZN-B707. (D) *L. hargravesii* SZN-B781. (E) *L. hargravesii*, strain SZN-B772. (F) *T. belgicus*, strain SZN-B739 and SZN-B755. (G) comparative peaks among the species.

A

m/z	R_t	λ_{max}	compound
***L. apeorus* SZN-B727**			
309.2	23.04	235	16-OH
381.2	22.07	234	OH 22:6
355.1	19.15	234	OH 20:5
371.1	14.32	210	EPA 20:5
371.2	13.09	210	EPA 20:5
301.1	12.62	-	
***L. aporus* SZN-B764**			
309.2	23.12	235	16-OH
381.2	22.14	234	OH 22:6
338.2	20.62	-	
355.1	19.15	234	OH 20:5
345.1	15.36	-	
371.2	14.26	210	EPA 20:5
301.1	12.58	-	
310.2	13.37	-	

B

m/z	R_t	λ_{max}	compound
***L. convexus* SZN-B783**			
309.2	23.04	235	16-OH
381.2	22.07	234	OH 22:6
355.1	19.12	234	OH 20:5
353.2	18.67	234	
397.2	17.15	210	EPA 22:6
395.2	15.91	234	
310.2	13.46	-	
371.1	14.05	210	EPA 20:5
301.1	12.62		
***L. convexus* SZN-B778**			
309.2	23.11	235	16-OH
381.2	22.08	234	OH 22:6
355.1	19.12	234	OH 20:5
353.1	18.66	234	
397.2	17.12	210	EPA 22:6
395.1	15.83	234	
371.1	13.99	210	EPA 20:5
371.1	12.49	210	EPA 20:5
371.2	12.12	210	EPA 20:5
301.1	11.40	-	

C

m/z	R_t	λ_{max}	compound
***L. danicus* SZN-B714**			
309.2	23.08	235	16-OH
381.2	22.28	234	OH 22:6
338.2	20.64	-	
355.2	19.96	234	
355.1	19.42	234	OH 20:5
397.2	17.71	210	EPA 20:5
371.1	14.66	210	
310.2	13.38	-	
301.1	12.59	-	
***L. danicus* SZN-B707**			
309.2	23.08	235	16-OH
381.2	22.28	234	OH 22:6
355.1	19.41	234	OH 20:5
397.2	17.72	210	EPA 22:6
371.1	14.68	210	EPA 20:5
371.2	13.41	210	
301.1	12.64	-	

D

m/z	R_t	λ_{max}	compound
***L. hargravesii* SZN-B781**			
309.2	23.03	235	16-OH
381.1	22.27	234	OH 22:6
355.2	21.09	234	
355.1	19.45	234	OH 20:5
331.2	19.07	-	
397.1	17.71	210	EPA 22:6
373.2	17.01	-	
397.2	15.93	210	
371.1	14.66	210	EPA 20:5
371.1	14.31	210	EPA 20:5
371.1	12.91	210	
301.1	12.59	-	
371.1	12.24	210	
301.1	11.23	-	

Contd...

Contd...

E

m/z	R_t	λ_{max}	compound
***L. hargravesii* SZN-B772**			
309.2	23.12	235	16-OH
381.1	22.31	234	OH 22:6
355.2	21.09	234	
355.1	19.48	234	OH 20:5
331.3	19.01	-	
397.1	17.67	210	EPA 22:6
373.2	16.99	-	
371.1	14.59	210	EPA 20:5
371.1	14.33	210	EPA 20:5
371.1	12.89	210	
313.1	12.06	-	
301.1	11.23	-	

F

m/z	R_t	λ_{max}	compound
***T. belgicus* SZN-B739**			
309.2	23.06	235	16-OH
397.2	22.23	210	EPA 22:6
355.2	21.06	234	OH 20:5
338.2	20.54	-	
395.2	18.8	234	
371.2	16.1	210	EPA 20:5
371.2	15.46	210	EPA 20:5
371.2	14.02	210	
310.2	13.4	-	
301.1	12.61	-	
***T. belgicus* SZN-B755**			
309.2	23.11	235	16-OH
397.2	22.28	210	EPA 22:6
355.1	21.11	234	OH 20:5
395.1	18.8	234	
371.2	16.07	210	EPA 20:5
371.2	15.45	210	EPA 20:5
371.2	13.95	210	EPA 20:5
301.1	12.58	-	

G

L. aporus						
L. convexus	353/18.6	395/15.9	397/17	355/19	381/22	371; 301/12
L. danicus						
L. hargravesii	331/19	373/17				
T. belgicus		395/18.8	397/22	355/21		

* mass/retention time

Legend:

16-OH= 16-hydroxyhexadecaenoic acid methyl ester as internal standard.
OH 20:5= Hydroxy acid of C20:5 fatty acid
OH 20:5= Hydroxy acid of C22:6 fatty acid
EPA 20:5= Epoxy alcohol of C20:5 fatty acid
EPA 20:5= Epoxy alcohol of C22:6 fatty acid

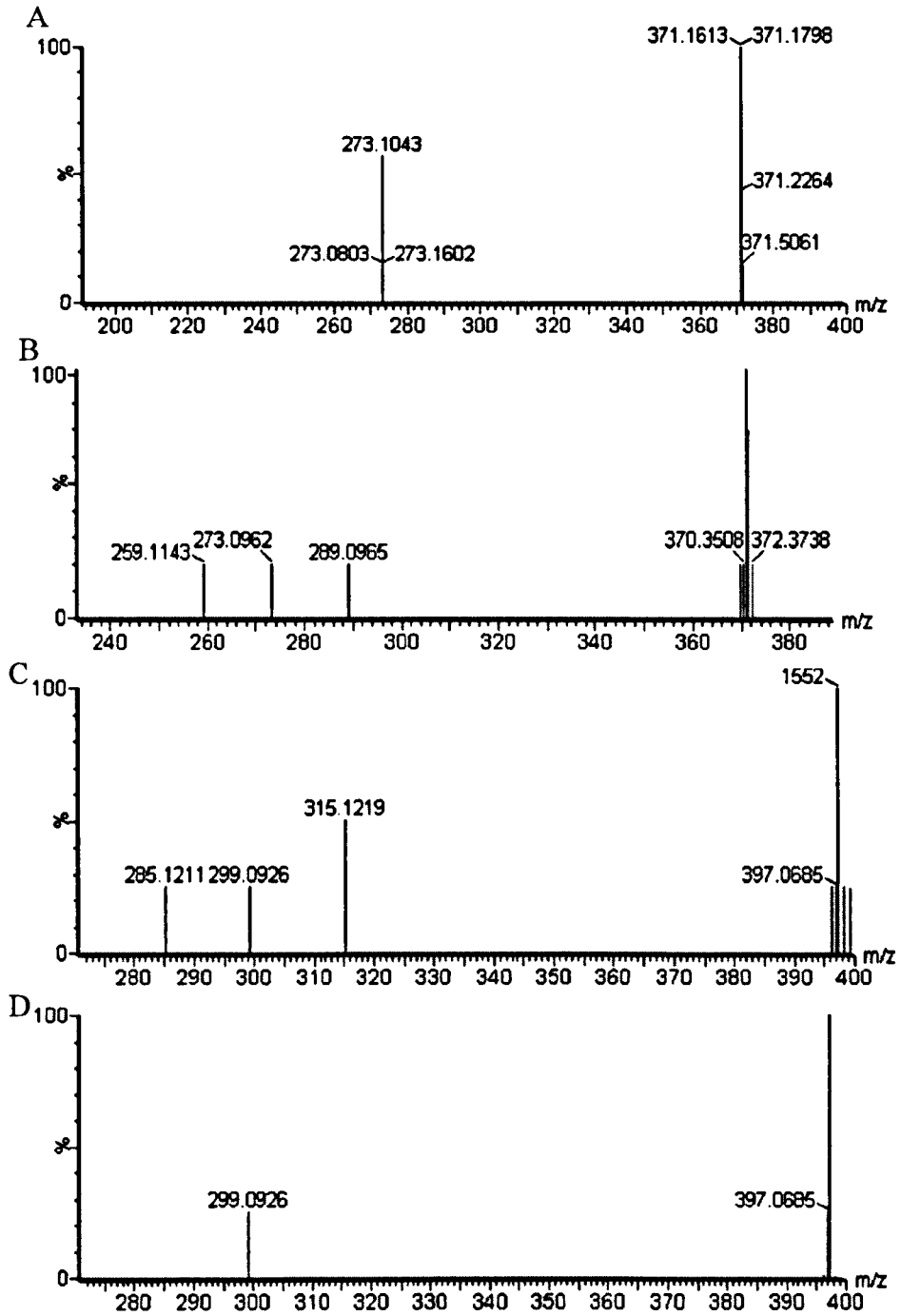


Fig. 4.4. MS/MS fragmentation of (A) 16-hydroxy-14,15-epoxy-eicosa 5Z,8Z,11Z,17Z-tetraenoic acid isolated from *L. aporus* and *L. convexus* (all strains). (B) 13-hydroxy-14,15-epoxyeicosa-5Z,8Z,11Z,17Z-tetraenoic acid isolated in *L. danicus* and *L. hargravesii* (all strains). (C) 18-hydroxy-16,17-epoxydocosa-4Z,7Z,10Z,13Z,19Z-pentaenoic acid isolated from *L . aporus* and *L. convexus* (all strains). (D) 15-hydroxy-16,17-epoxydocosa-4Z,7Z,10Z,13Z,19Z-pentaenoic acid isolated from *L. danicus* and *L. hargravesii* (all strains).

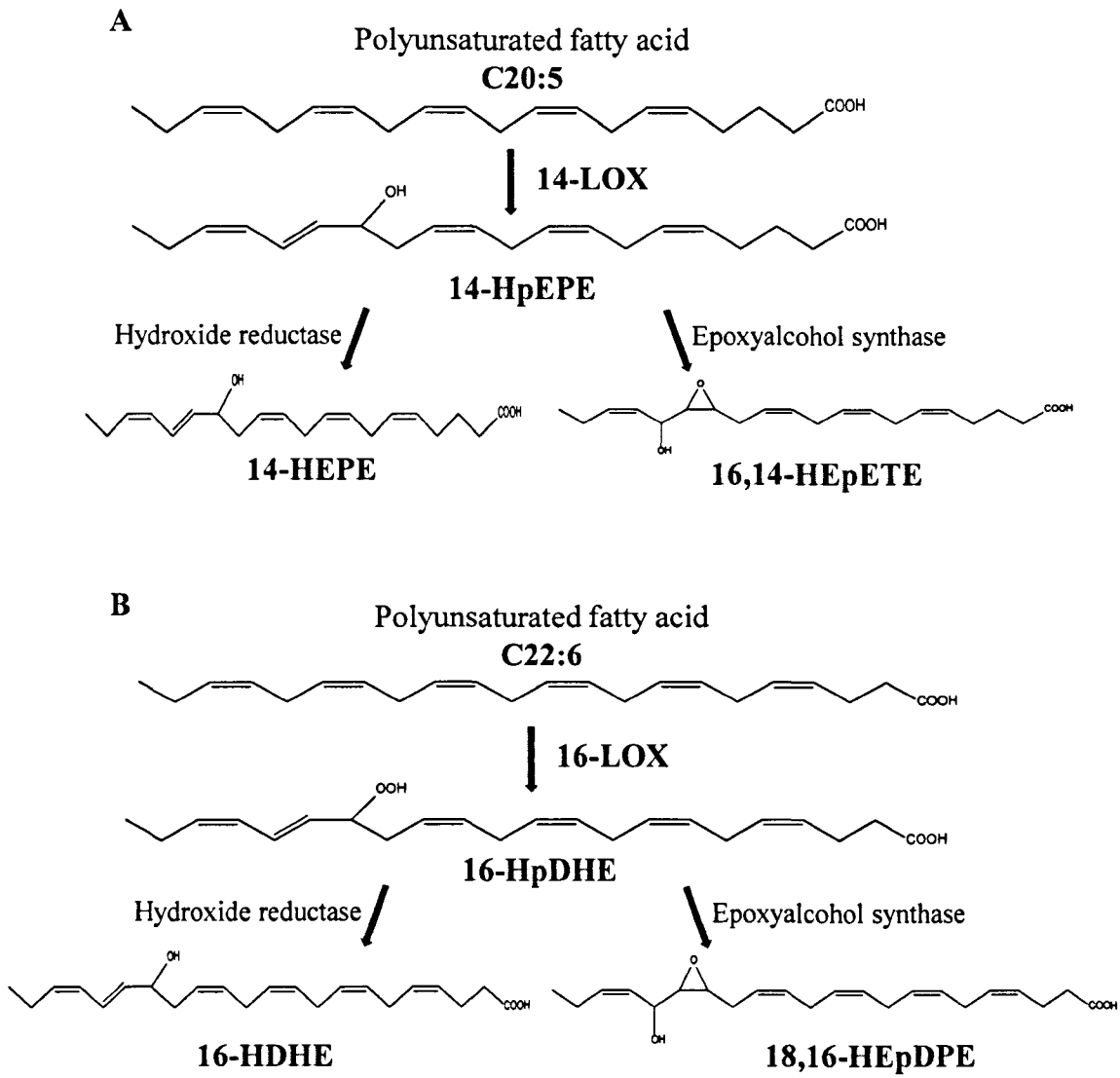
The MS/MS spectra of epoxyalcohol with m/z 371, R_t 14 & λ_{max} 210 of *L. danicus* and *L. hargravesii* (Table 4.2, C–E) showed fragments at m/z 289, 273 and 259 accounting for the occurrence of the intermediate compound containing the a-hydroxy-aldehyde function of 13-hydroxy-14-oxo-5,8,11-tetradecatrienoic acid methyl ester (Fig. 4.4, B). These ions were consistent with the structure of 13-hydroxy-14,15-epoxyeicosa-5Z,8Z,11Z,17Z-tetraenoic acid (13,14-HEpETE).

The analysis of MS/MS spectra of epoxyalcohol with m/z 397, R_t observed in *L. aporus* and *L. convexus* revealed the presence of three fragments of m/z 285, 299 and 315, that are consistent with molecule 15-hydroxy-16,17-epoxydocosa-4Z,7Z,10Z,13Z,19Z-pentaenoic acid (15,16-HEpDPE; Fig. 4.4, D).

The MS/MS spectrum of epoxyalcohol with m/z 397, R_t 17 recorded in *L. aporus* and *L. convexus* (Table 4.2, A and B) showed a single molecular ion at m/z 299, that was diagnostic to assign to the structure of 18-hydroxy-16,17-epoxydocosa-4Z,7Z,10Z,13Z,19Z-pentaenoic acid (18,16-HEpDPE; Fig. 4.4, D) .

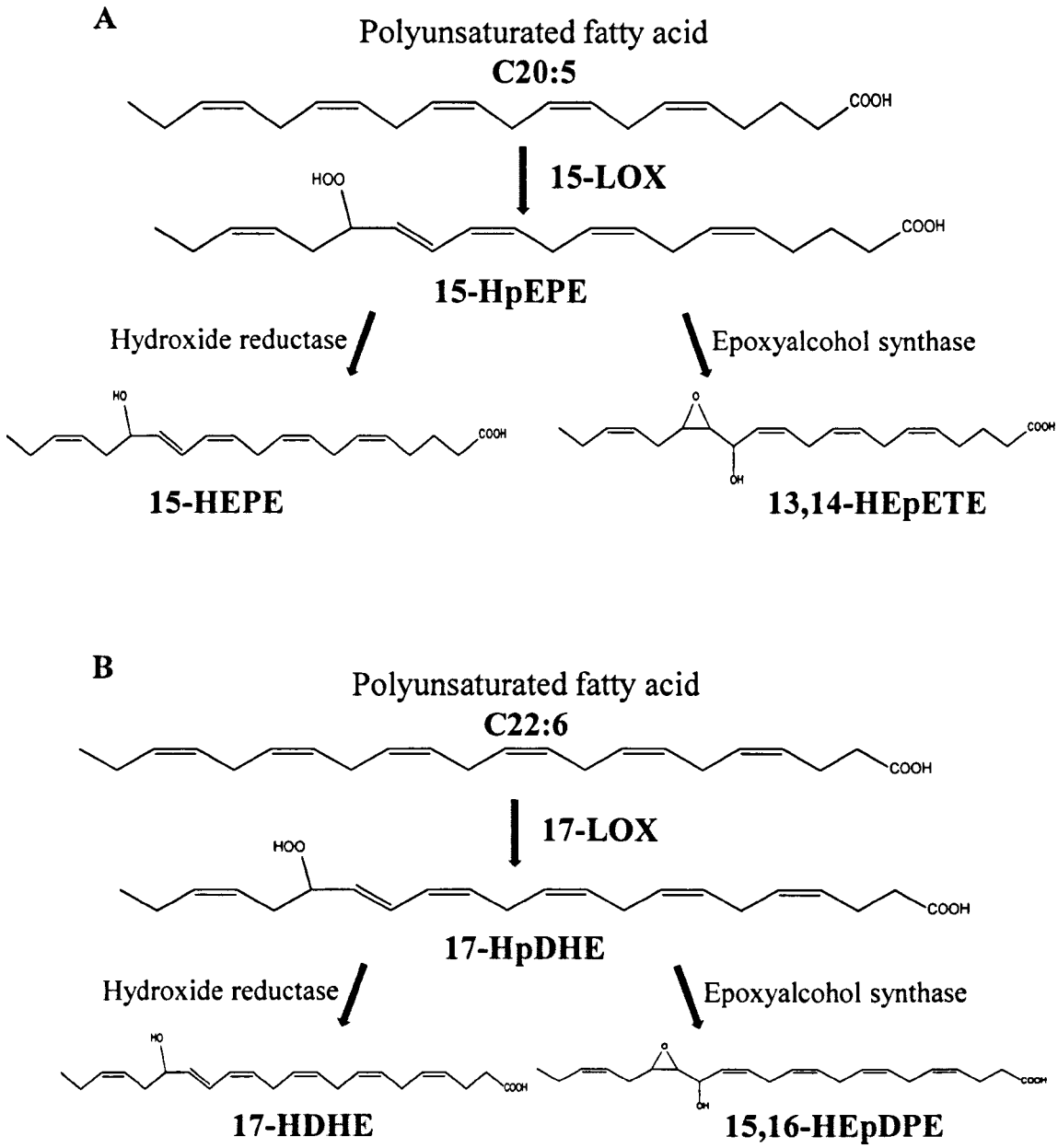
The compounds with m/z 353 and 395 remained uncharacterized, because they had m/z and R_t that are not diagnostic for their characterization with existing knowledge of these compounds. Isolation and structural elucidation is necessary to assign a structure to these molecules, which are planned for future studies.

Likewise, on the basis of the epoxyalcohol found in each species, lipoxygenase pathways were established. In *L. aporus* and *L. convexus,* 14-LOX acting on EPA and 16-LOX specific for DHA were identified (Fig. 4.5, A and B). Instead, in *L. danicus* and *L. hargravesii* 15-LOX activity that recognize EPA and 17-LOX acting on DHA was characterised (Fig. 4.6, A and B).



14-HpEPE = 14-Hydroperoxy eicosapentaenoic acid
14-HEPE = 14S- Hydroperoxy eicosa-5Z,8Z,11Z,13E-triaenoic acid
16,14-HEpETE = 16-hydroxy-14,15-epoxy-eicosa 5Z,8Z,11Z,17Z-tetraenoic acid
16-HpEPE = 16-Hydroperoxy eicosapentaenoic acid
16-HDHE = 16-Hydroxy docosa hexaenoic acid
16-HEpDHPE = 18-hydroxy-16,17-epoxydocosa-4Z,7Z,10Z,13Z,19Z-pentaenoic acid

Fig. 4.5. Oxylipins synthesis pathways of C20:5 and C22:6 PUFAs in *L. aporus* and *L. convexus*. (A) 14-LOX pathway of EPA. (B) 16-LOX pathway of DHA.



15-HpEPE = 15-Hydroperoxy eicosapentaenoic acid
15-HEPE = 15S- Hydroperoxy eicosa-5Z,8Z,11Z,13E-triaenoic acid
13,14-HEpETE = 13,14-threo-13R-Hydroxy-14S,15S-trans-epoxyeicosa-5Z,8Z,11Z,13E-triaenoic acid
17-HpEPE = 17-Hydroperoxy eicosapentaenoic acid
17-HDHE = 17-Hydroxy docosa hexaenoic acid
17-HEpDHPE = 17-Hydroxy epoxy docosa pentaenoic acid

Fig. 4.6. Oxylipins synthesis pathways of C20:5 and C22:6 PUFAs in *L. danicus* and *L. hargravesii.* (A) 15-LOX pathway of EPA. (B) 17-LOX pathway of DHA.

4.4. Discussion

In Chapter II, through the application of a combination of molecular makers from nuclear and plastid origin and morphological characters the species of *Leptocylindrus* and *T. belgicus* were distinguished and their relationships are inferred. In the light of those results in the present study, differences among the newly recognised species in the synthetic pathways of oxylipin are analysed. The study of oxylipin composition and possible biosynthetic pathways in the already known and newly recognised species offers a new angle for the study of their diversity. In the following, the results obtained are discussed in the relationships evidenced among the species in Chapter II.

4.4.1. ***Support for the species recognized in Chapter II.*** The differences noted in the growth curves of species suggests that the three lineages identified in Chapter II have some distinguishing feature in their growth which reflect their phylogenetic distance. The analysed oxylipin pathways of *Leptocylindrus* species and *T. belgicus* indicate that the hydroxyacids and epoxyalcohol of EPA and DHA are the common traits of these genera. The LOX pathways between the species lineages (Chapter II), *L. aporus* and *L. convexus* (Lineage II; Fig. 4.5, A and B), were different from those found in *L. danicus* and *L. hargravesii* (Lineage III; Fig. 4.6, A and B) indicating that the two lineages have evolved different LOX for oxylipin synthesis. Among the species, *L. danicus* and *L. hargravesii,* and *L. convexus* and *L. aporus,* showed very similar oxylipin profiles, in which hydroxyacids and epoxyalcohol of EPA and DHA had the same R_t. In fact, Lineage II is characterised by 14-LOX acting on EPA and 16-LOX acting for DHA, whereas in the Lineage III is characterized by 15-LOX activity that recognizes EPA and 17-LOX activity that recognizes DHA. In comparison with Lineage II and III constituted by *Leptocylindrus* species, *Tenuicylindrus belgicus* (Lineage I) has quite dissimilar oxylipin pathways,

showing mainly the presence of hydroxyacid of EPA, and the complete absence of DHA derivatives. The sharing of common pathways across Lineage II and Lineage III reflects close association of the species within the two lineages, whereas their distance from Lineage I validate the molecular, morphological and phylogenetic relationships delineated in Chapter II.

In addition to the compounds identified and characterised as lineage specific other peaks observed in each of the species can be further investigated to provide a species specific compound pattern. For instance the compound with m/z 395 (R_t 15.9, λ_{max} 253) was specific to the species *L. convexus,* hence can be used to recognise *L. convexus* among the other *Leptocylindrus* species. Nevertheless, here the lineage specific LOX pathways are identified. Therefore in case of *Leptocylindrus* and *Tenuicylindrus* oxylipins are potential biomarkers.

4.4.2. ***Diatom taxonomic classification.*** Diatoms taxonomy has traditionally relied on the differences in the frustule structure for species classification and delimitation. However, with the discovery of PCR and DNA sequencing, molecular markers have emerged as the more reliable markers for species discrimination. Despite the availability of multiple molecular markers the scientific community is still to agree upon a consensus gene for the barcode of algae and plants, like that of *cox* gene in animal system. A combination of rbcL and LSU are proposed, rbcL is proposed to be used as the primary marker for diatom barcoding, while LSU D2/D3 to be sequenced as a secondary marker to facilitate environmental surveys (Hamsher *et al.* 2011).

4.4.3. ***Chemotaxonomy in Diatoms.*** Biochemical fingerprinting is not very new in species identification, but rather has decades of history. In fact, allozymes (principle - protein variants in enzymes that can be distinguished by native gel electrophoresis according to

differences in size and charge caused by amino acid substitutions) were the first molecular markers to be established (Schlotterer 2004). These markers showed high amounts of within-population polymorphism that led to a very important discovery, the neutral theory of evolution. The theory essentially states that most of the mutations are effectively neutral (Kimura 1968, King and Jukes 1969). However due to scarcely informative marker loci, the use of allozymes for mapping and association studies (i.e. the joint occurrence of two genetically encoded characteristics in a population and, often, an association between a genetic marker and a phenotype) was challenged (Hills *et al.* 1996). In recent years, with the advancement of technologies, biochemical fingerprinting has emerged as a valuable alternative. In diatoms this approach, when applied to the discrimination of closely related species, could also highlight physiological differences and hence provide valuable information on the ecological characteristics of the distinct species.

Other molecules, such as pigments (Ruivo *et al.* 2011) and PUFAs or in general hydrocarbons, are used for species discrimination in general and also in diatoms. An example is the cuticular hydrocarbons in insects (Singer 1998, Baracchi *et al.* 2010). In phytoplankton, pigment concentrations and ratios are investigated for identifying species or lineage specific character. Irradiance-independent, species- or group-specific discrimination was recently demonstrated on 18 species of *Pseudo-nitzschia* spp., a genus that has a complex genetic structure (Zapata *et al.* 2011). Comparative pigment (chlorophylls and carotenoids) concentrations studied in cultures of dinophyceae, bacillariophyceae, prasinophyceae and cryptophyceae showed that growth phase and irradiance can affect the levels and ratios of pigments in cells (Ruivo *et al.* 2011). Nevertheless, pigments are being investigated for the applicability in detection of dominant groups based on the ratios in natural environments (Ruivo *et al.* 2011).

Many studies have investigated the hydrocarbons or PUFAs composition of microalgae, but the characterisation of these compounds is largely limited to the exploration of long-chain PUFAs for commercial production (Volkman *et al.* 1989), with few addressing their application for taxonomy (Lang *et al.* 2011). Fatty acid profiles of 2976 strains of microalgae (from different algal groups) were suitable to discriminate taxa of higher rank, while the variability was too high to obtain a species-specific character (Dunstan *et al.* 1993). Thus a systematic use of these compounds in diatom taxonomy awaits investigation.

4.4.4. ***Oxylipins as biomarkers.*** Since the discovery of the diatom PUAs role as a defence mechanism against grazers, many studies have also addressed (focused on) the other possible functional roles of these molecules. Such studies, although few, have contributed to the understanding that these molecules are highly diverse, at times lineage-specific, and are involved in complex biological interactions. In the current study, using targeted HPLC coupled with LC-MS oxylipin profiling of *Leptocylindrus* species and *T. belgicus*, a unique profile for each species was generated that allows in species identification. The study demonstrates the diversity in the synthetic pathways of oxylipins and the possible application of these compounds as biomarkers for species delimitation.

Oxylipins have rarely been applied in species delimitation due to the plasticity in the production of these compounds. Gerecht et al. (2011) revealed that oxylipin production in *Skeletonema marinoi* can vary even among clones isolated from the same area in different time periods. These authors reported significantly higher amount of oxylipin in two of the four strains analysed. In addition, qualitative differences in oxylipin production among the clones were noted that might provide competitive advantages. This result was also consistent with the genetic diversity observed among the clones of the species that might further have propagated downstream (Gerecht *et al.* 2011). Thus it appears that the diversity of oxylipins can even be traced back to intraspecific genetic differences. The

clones used in the current study are from two different time periods but during the same season. Moreover, the differences described among the species are consistent between the strains, while quantitative differences are excluded for species and lineage delineation. Considering that the small genetic differences in molecular marker genes observed in the clones of the species *S. marinoi* was replicated in the oxylipin content of the cells (Gerecht *et al.* 2011), it seems obvious that different taxa might have diverse oxylipins deriving from diverse enzymatic activity. The genetic relationships constructed with the phylogenetic marker genes in this thesis (described in Chapter II) are consistent with the biochemical synthetic pathways observed for the oxylipins. *Leptocylindrus aporus* and *L. convexus* that belong to same lineage in the phylogenetic tree (Chapter II on this thesis) had similar oxylipin pathways. The two pathways of *L. danicus* and *L. hargravesii* are highly similar indicating close relationship. This was also revealed in the molecular analysis where the species could be differentiated with the rbcL, psbC and ITS regions but not with LSU and SSU genes. *Tenuicylindrus belgicus*, a species for which a separate genus has been established (Chapter II) based on remarkable structural-ultrastructural and molecular differences, also shows conspicuous differences in oxylipin synthesis pathways and the compounds produced within. Similar results have also been reported in another study involving sympatric species of *Pseudo-nitzschia delicatissima,* where the reproductively isolated pseudo-cryptic species (Amato *et al.* 2007) were also qualitatively and quantitatively differentiated in LOX products (Lamari 2009). Thus greater efforts in characterising the pathways of oxylipins synthesis in diverse diatoms can help in understanding its applicability as a potential biomarker.

4.4.5. *Intraspecific variations in oxylipins.* During growth progression along the growth curve and bloom time obvious metabolic changes occur (Vidoudez and Pohnert 2011). Changes have also been noted in the production of oxylipins, as reported in some recent

studies (Vidoudez and Pohnert 2008, d'Ippolito *et al.* 2009). In the current studies, oxylipins were extracted from the stationary phase cultures, since a significantly higher oxylipins production was reported during this phase in other species including *S. marinoi* (Ribalet *et al.* 2007, Vidoudez and Pohnert 2008) and *P. delicatissima* (d'Ippolito *et al.* 2009). Therefore oxylipin machinery is functional at its maximum during the stationary phase.

Environment exerts its influence on all the biological process and so does on the production of oxylipins in diatom cells. Physiological stress has been reported to trigger the production of oxylipins. For example, N and P limitation in *S. marinoi* increased PUAs production by 1.4 to 1.8 fold (Ribalet *et al.* 2007). Silica- limited cells of *S. marinoi* have been shown to produce a 7.5 fold higher amount of PUAs (Ribalet *et al.* 2009). Additionally, temperature has also been reported to have a negative relationship with EPA content in *Nanochloropsis* sp. (Hu and Gao 2006). This relationship with temperature has been discussed in more detail in Chapter V. Consequently, environmental factors particularly nutrient availability can profoundly influence the oxylipin content in diatoms. Nevertheless in the study all the environmental factors were kept constant between the species.

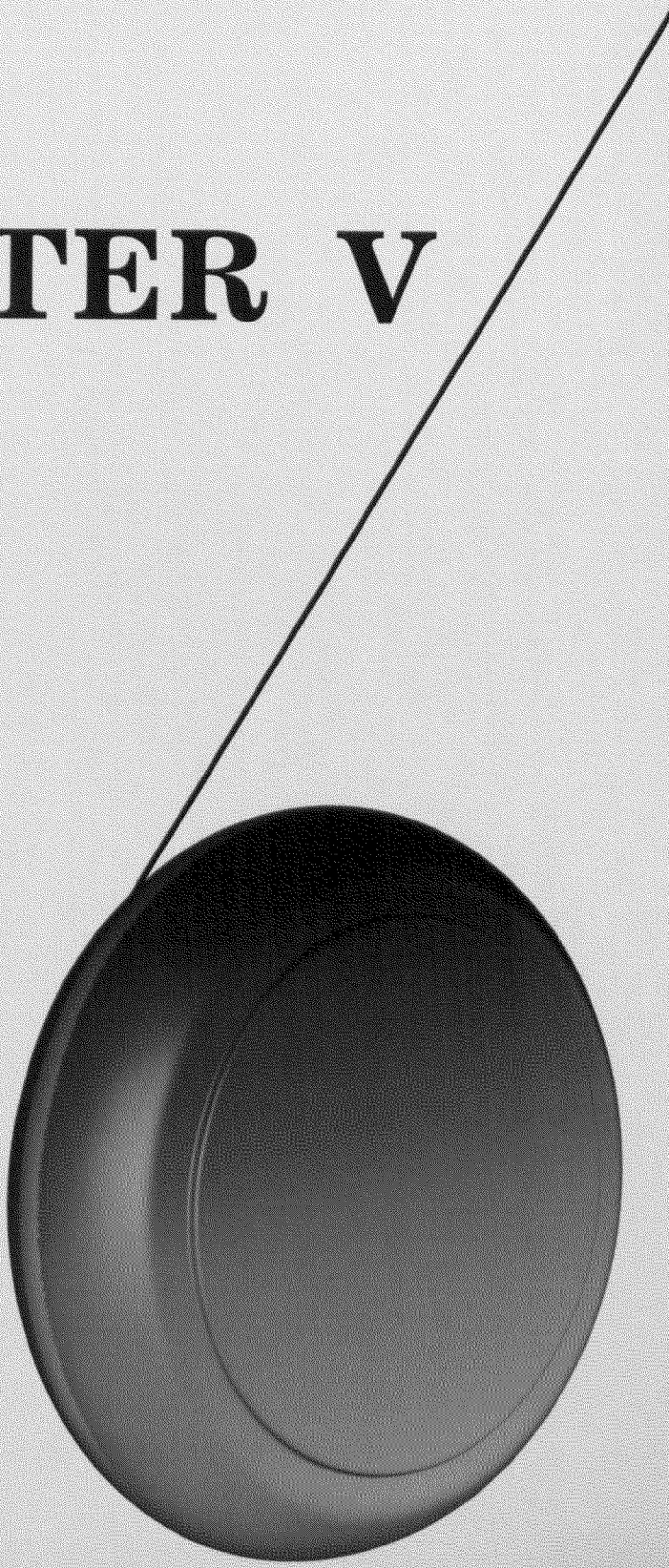
4.5 Conclusion

In the current study, a method for the identification of individual *Leptocylindrus* species and *T. belgicus* using the oxylipin profiles is proposed. A methodology is also proposed as a metabolic discrimination of the species identified using the morphological and DNA-based molecular markers. The application of morphological, molecular and metabolic discrimination of species will provide a fertile interface to diatom species identification and physiological validation of the different genetic entities that might have different ecological roles.

Oxylipin profiling for species delimitation has several practical advantages that are tantalizing to employ them as biochemical markers either as sole markers or as a support to morphologically and physiologically discriminated genetic entities. The protocol described is high throughput, relatively economic on a per-sample basis and can be fully automated with analytical acquisition time taking less than three hour. PUFAs account for a large portion of the diatom cell volume and therefore only a few cells are needed for the analysis. Additionally, the use of internal standards, such as 16-hydroxyhexaenoic acid in the analysis, allows for the quantitative comparison of the phyco-oxylipin fingerprints among species.

The diversity in oxylipins class appears to be too high and certainly provides enough characters for discrimination even for the species rich diatoms or even a combination of different compounds can reflect the species identity. On the other hand, the process of oxylipin synthesis is dynamic and is influenced by the biotic and abiotic factors (discussed in Chapter V). Therefore, prior to its application as biomarker, it is important to decipher the factors that trigger and the physiological condition that favour their synthesis. Thus with the greater efforts in characterising the pathways of oxylipins synthesis in diverse diatoms species can help not just in understanding its applicability as a potential biomarker but also shed light on numerous other process in which these molecules are involved.

CHAPTER V



Physiological responses and biochemical phenotypes in the diatom species *Leptocylindrus danicus* and *Leptocylindrus aporus* exposed to three different temperatures

5.1 Introduction

One of the intriguing questions in the study of phytoplankton ecology is how these microalgae have been able to adapt to the extremely wide spectrum of environmental conditions. Since their origin (Triassic-Jurassic period), microalgae of the contemporary ocean have been constantly adapting to the ever changing marine environment and this fight against the dynamic nature of the oceans has been conducted by increasing fitness and producing diverse genotypes to suit either narrow or diverse niches.

It is generally accepted that diatoms have a wide thermal tolerance agreeing with Eppley's (1972) that 'temperature does not seem to be very important in the production of phytoplankton in the sea' and that maximum growth rate of phytoplankton scale with temperature (Eppley 1972). Thus they have successfully colonised diverse spatial and temporal thermal environments. Although diatoms are ubiquitous as a group, their diversity is wide. For example their morphology ranges from large, chain forming species, such as *Fragilariopsis kerguelensis* found in polar latitudes (Nodder and Waite 2001), to small unicellular species in tropical waters (Cavender-Bares *et al.* 1999). The cosmopolitan species, which have a wider geographical range including diverse environmental conditions, often show considerable ecological plasticity (Smayda 1958, Hasle 1976). In contrast, there are also species which have a limited tolerance to environmental variations and hence are endemic to a restricted geographical location. For example, a careful morphological analysis of the Thalassonemataceae species revealed that, of the few species occurring in temperate to cold waters, one has a cosmopolitan distribution, one is restricted to the northern hemisphere and two others to the Southern Ocean (Hasle 2001). These studies reveal that diatom distribution is influenced by environmental conditions in general

and temperature in particular and there is a necessity to understand this influence on the growth and physiology of individual species.

Among the different environmental factors which the cells cope with, temperature plays a major role driving the spatial and temporal distribution of a species, and thus the community composition, and food web structures (Pörtner and Farrell 2008, Montes-Hugo *et al.* 2009). Temperature is one of the fundamental factors because growth is only possible within some relatively limited range. In some groups temperature plays a key role for example, the prochlorophyte, *Prochlorococcus* cannot grow below 15–18°C (Olson *et al.* 1990, Cavender-Bares *et al.* 2001). In other cases, temperature plays a secondary role i.e. when coupled with other environmental controls such as nutrient availability and light, having either an antagonistic or a synergistic role in structuring the species distribution. For instance, a marked increase in diatom abundance was noted when both Fe supply and temperature were increased (Rose *et al.* 2009). In diatoms, where the effect or role of temperature is still unresolved in most cases, findings from laboratory and/or field experiments combined with distribution data can help understanding the influence of temperature on individual species distribution.

Temperature can have substantial effect on the growth and physiology of the cells by altering the rate of enzyme catalysed reactions in various metabolic processes, changing the cell composition. Temperature also influences many physical cellular processes such as diffusivity, solubility and fluidity that are associated with growth, which are maintained in a feedback control processes with the external environment. These effects are species and at times strain specific reactions and they determine the successes in coping with the fluctuating temperature. It is important to understand these differences if it is possible to determine the significance of the physiological variations. They are primarily reflected in the physiological adjustments of cellular mechanisms and finally in the biochemical phenotype i.e. the metabolome of cells.

The metabolome is the final product of genetic and environmental interactions, and as such provides a blueprint for the cellular biochemical machinery. Understanding the regulation of the metabolome enables a better understanding of the biochemical pathways that control biological processes.

Metabolomics, i.e. the identification and quantification of all the metabolites in a biological samples (Dettmer *et al.* 2007), has been successfully applied in studies of environmental influence on cells, for example, stress physiology under nutrient depletion (Bölling and Fiehn 2005), acclimation to CO_2 limitation (Renberg *et al.* 2010) in *Chlamydomonas reinhardtii*, and other ecological and evolutionary studies (Macel *et al.* 2010). Although metabolomic studies related to microalgae are limited, (Bölling and Fiehn 2005) they have already provided significant insights in to molecular mechanisms and have identified novel metabolites and pathways that are triggered by changes in the environmental condition.

Diatoms possess a wide array of biochemical compounds that help them mitigating the effects of a changing environment and often these compounds are released into the environment (Hay 1996), which may provide an added advantage such as defence against grazers (ex. Domoic acid, Shaw 1997; Saxitoxin in dinoflagellates, Schantz, 1966 and pathogens, Naviner 1999) or in overcoming competition for resources or to overcome stress conditions. These compounds, referred to as secondary metabolites, are low molecular weight molecules with diverse chemical structures, among which are the physiologically and ecological relevant group of fatty acid derivatives called oxylipins (Chapter IV). Secondary metabolites including oxylipins are often modified in terms of their concentrations or functional properties to help make the organism more fit to its environment and at times can be used as signature molecules that reflect the physiological status and affect response to stress (d'Ippolito *et al.* 2004, Vidoudez and Pohnert 2008, Vidoudez and Pohnert 2011). Despite all these implications, the research of diatom

metabolites has a very short history, with limited knowledge on chemical diversity and exploration of biosynthetic mechanisms with the primary goal of isolating novel compounds in the search for new drugs.

Response to temperature has been well studied in the terrestrial plant system, at least in comparison to microalgae, through the application of system biology approaches involving a wide array of genetic, molecular and metabolomic techniques. Non-targeted metabolome analyses of low temperature acclimation have revealed that the plants reconfigure the metabolome extensively. The central carbohydrate metabolism was found to be most strongly positively correlated metabolites, increasing the levels of glucose, fructose, and sucrose (Hannah 2006). Another study reported an increase in the metabolites like trehalose, putrescine and ascorbate, all of which are considered to have potential roles in cold tolerance and a possible up regulation in the urea cycle (Cook *et al.* 2004). High temperature metabolome response included an increase in the level of beneficial osmolytes including sucrose, maltose, trehalose, fructose and glucose. At high temperature proline was not found to be beneficial, since it was down regulated, whereas glutamine synthesis was specifically elevated. Thus, in plants, in response to temperature stress at high and low temperature, there is a profound shift in the levels of amino acids, citric acid cycle intermediates and many other metabolites of carbohydrate metabolism.

In diatoms, the response to temperature stress is mainly addressed in studies where this factor is considered as synergistic/antagonist in combination with other factors like nutrient limitation, salinity. Temperature stress metabolome in diatoms still needs to be explored and studied, it might or might not respond in the same pattern as other photosynthetic organisms. For example, the response of *Thalassiosira pseudonana* carbon metabolism to nitrogen starvation was found to be similar to cyanobacteria but not to green alga and *Arabidopsis* (Hannah *et al.* 2010). In general diatom species physiologically acclimate in response to variation in temperature, for example, by adjusting their membrane lipid

composition, thus increasing their capacity to grow. Some of the responses reported include an increased level of unsaturated fatty acids as a marker of low temperature adaptation in *Chaetoceros socialis* (Degerlund *et al.* 2012), and an increase in the level of proline and chlorophyll content per cell was noted for cold region strain than the warm region strain in *Phaeodactylum tricornutum* as a response to increase in temperature (Kräbs and Büchel 2011). Another study in *Chaetoceros* cf. *wighamii* revealed higher carbohydrates and lipid content at lower temperatures than at higher temperature tested, whereas there was no significant change in the protein levels (de Castro Araújo and Garcia 2005). Thus it appears diatoms respond to temperature in concordance with the higher plant system, but the hypothesis needs to further tested and studied.

To investigate whether the diversity found in the morphology and phylogenetics is paired with similar degree of diversification at the physiological and biochemical level, differences in growth characteristics and metabolite levels at varying temperature were analysed in related species. Ultimately, the study could also help to understand how biochemical constitutes are linked to the fitness of the species in relation to the particular suite of environmental or ecological conditions it faces in nature.

The study focuses on the ecophysiology of *Leptocylindrus danicus* and *L. aporus*, which were selected among the other *Leptocylindrus* species (Chapter II) due to their relevance in the Gulf of Naples (GoN) phytoplankton. The two species, *L. danicus* and *L. aporus*, have been strategically chosen based on their remarkable differences in the phylogeny and time of occurrence in the natural environment, despite subtle morphological differences. We believe the system is simple as well as robust enough to understand the problem. At the sampling site LTER-MareChiara, typically, water temperature varies between 12 °C and 22 °C from winter through spring, while during summer it is between 22 °C and 26 °C (Ribera d'Alcalà *et al.* 2004). The contrasting temperature conditions associated with the occurrence of the two species (Chapter II) suggest they might have opposite physiological

attributes that are suitable to the environment in which they occur. Thus, the aim of the experiment was to establish the temperature tolerances of the two species and to understand the biochemical phenotype specific to the temperature and species. To this end, cultures of species were exposed to three temperature conditions of 12 °C, 19 °C and 26 °C. During the experiments,

1) growth rates, as net products of sub-cellular responses (Munns and Tester 2008),

2) maximum cell density, indicating the carrying capacity of the species,

3) cell morphology, as a reflection of the vigour of cultures, and

4) selected metabolites, as the clues of physiological adjustments, were analysed. Based on their occurrence, the response of the two species at lower temperatures and higher temperature might be different. Furthermore, the biochemical phenotype though whole metabolite, that are planned for later studies, might help in understanding the molecular mechanisms that are altered as a response to variations in temperature.

The study included an acclimation step, in which four strains of each species were gradually adapted to the different temperature conditions, followed by an analysis of their acclimated growth response and of possible intraspecific variations of growth parameters. One strain for each species was then selected and submitted to new growth experiments at the three temperatures prior to oxylipins and fatty acid analyses. The latter analyses were conducted in collaboration with Dr Angelo Fontana and colleagues at CNR-*Instituto di Chimica Biomolecolare*, Naples, Italy, where the facilities of LC-MS and GC-MS were utilised, together with scientific guidance. Overall, results of this study support the idea that functional variations in diatoms may or may not be reflected in profound or subtle differences in the external silica frustules. Therefore, it is proposed that, at least in this case, species physiological properties could be a better indicator of phylogenetic relationships than morphology. The study also restates the importance of addressing

genetic variations among similar species, as they could be indicator of physiological differences that are relevant to the understanding of the ecology of these species.

5.2 Methods:

5.2.1 Cultures. Four strains of *L. danicus* and *L. aporus* were selected (Table 5.1) out of those used for the study of genetic diversity at the GoN (Chapter II). The eight strains were chosen among those isolated on different dates, to cover the entire season of occurrence of the two species and possibly identify intraspecific variations in the adaptation to different environmental conditions. Two strains of *L. danicus* and *L. minimus* submitted to the Moore Foundations for transcriptome analysis were included among the selected strains. Stock cultures were maintained in an incubator at 20 °C and under a 12:12 h L:D photoperiod, with an irradiance of about 40–60 µmol photons $m^{-2}sec^{-1}$. Cultures were maintained in the exponential phase by periodic transfers into 50 ml polystyrene flasks filled with 30 ml of K culture medium (Keller *et al.* 1987).

Table 5.1. Strains used in the physiological responses to three different conditions

Strain Information	Collection Date
Leptocylindrus danicus	
SZN-B707	15/02/2010
SZN-B714	21/04/2010
SZN-B715	15/06/2010
SZN-B650♠	15/06/2010
Leptocylindrus danicus	
SZN-B727	03/08/2010
SZN-B651♠	21/08/2010
SZN-B752	19/10/2010
SZN-B753	19/10/2010

♠ Strain used for whole transcriptome sequencing and metabolomic study

5.2.2 Acclimation, batch cultures and experimental conditions

i. Acclimation. Strains of both species were initially acclimated to experimental temperature (T) conditions in a 50 ml flask with 30 ml K media. Other factors including light intensity (100 µmol photons $m^{-2}sec^{-1}$), photoperiod (L:D, 12:12) and culture condition

(type of media and nutrient concentrations) were kept constant. For acclimation, cultures from the mother stock usually maintained at 19–21 °C were divided into three parts and incubated at T stepwise reduced or increased gradually reaching the planned experimental temperatures. To acclimate to lower T the cultures were transferred from 20 °C to 12 °C through 16 °C and 14 °C. To acclimate to higher temperature, the cultures were transferred from 20 °C to 26 °C through 22 °C and 24 °C. Throughout the acclimation period the cultures were maintained in the exponential phase by continuous dilution and, at each transition stage, *in-vivo* fluorescence was measured using a Turner 10-AU fluorometer, to estimate the cultured cell biomass. The cultures were grown at the different temperatures for about 25 to 35 days, depending on the experimental T. During the acclimation period cell diameter was measured cell to check whether there were variations that would indicate sexual reproduction or vegetative auto-enlargement. All cell measurements were made using an Axiophot light microscope (Carl Zeiss, Oberkochen, Germany) fitted with an ocular micrometre.

Subsequently to test acclimation, 200 ml flat polystyrene bottles containing 100 ml of K culture medium were inoculated with 1 ml of each culture in the exponential phase at a density of 2000 cells ml^{-1} until the stationary phase. At stationary phase, the cultures were diluted back to the initial density and allowed to grow again. Growth rate and *in-vivo* fluorescence, as growth indicators, were measured by taking daily samples of cultures starting from day two till the beginning of the stationary phase. A subsample of 4 ml was taken from each sample and cell counts were performed on 1 ml in a Sedgwick-Rafter counting chambers, preserving the remaining 3 ml culture for later verifications. Depending on the cell density, cell counts were performed in replicates of either transects or fields of counting chamber. *In-vivo* fluorescence was measured on a 15 ml subsample of experimental culture placed in a sterile glass tube, which was replaced back into the

respective culture bottles after the measurement. The whole procedure was repeated until the growth rate remained persistent for three sequential growth curves.

ii. Growth experiment with multiple strains. During the growth experiment, 150 cm^2 polystyrene flasks were inoculated with exponential cultures of the 8 strains selected (Table 5.1) to a density of 2000 cells ml^{-1}. The culture was allowed to grow at the respective T under 100 µmol photons $m^{-2}sec^{-1}$ of light and 12:12 L:D cycle. Each experiment typically lasted eight days and consisted of daily monitoring of the growth starting from the second day of inoculation. Cell density and *in vivo* fluorescence were monitored as described earlier (in acclimation).

iii. Growth and metabolomic experiment with two strains. during the metabolomic experiment, the strains of *L. danicus* and *L. aporus* also used for the transcriptome sequencing at the Moore Foundation (see above) were used to further study the physiological responses and to identify changes in selected biochemical properties under different T conditions. For each species, cells were inoculated from exponential growing cultures into five, 2 l spherical, flat bottomed, glass flasks with 1.2 l of K media to a cell density of 2000 cells ml^{-1}. The cultures were allowed to grow under 100 µmol photons $m^{-2}sec^{-1}$ with a 12:12 L:D cycle. All subsequent measurements were performed independently on each of the five flasks (i.e., at each T, for each species, n=5). Growth rate was monitored only by counting cell number per ml as described above for the growth experiment on the 8 strains. The changes in the morphological features, including variation in cell size, and colony formation were recorded at the same time in the light microscope. At the stationary or end of the exponential phase the cultures were centrifuged and the culture pellets were frozen until analysis. A culture pellet of 100 ml for protein analysis, 200 ml for oxylipin and polyunsaturated fatty acid (PUFA) analysis and 800 ml for whole metabolite profiling through NMR (for future studies) were prepared and frozen until used, for each of the 5 replicates of the two species.

5.2.3 Calculation of maximum growth rate, biovolume and biomass. To calculate maximum growth rates (μ_{max} = divisions day^{-1} or d^{-1}), cell density values were transformed into log_{10} values and analysed calculating the linear regression of log transformed cell density versus time for each species/replicate. The average and standard deviation were calculated.

Biovolume acts as a proxy for the physiological status of cells at different T values. Biovolume calculations were made using cell diameter measured on 20 randomly selected exponential phase cells along with the average pervalvar axis, obtained for each of the two species based on average values obtained in the later experiment with a single strain. To obtain more precise data on biomass achieved under different condition in the growth experiments on two strains with replicates, measurements of the diameter and pervalvar axis were made on 25 randomly selected cells. A standard cylindrical geometric shape for the cells was assumed to calculate biovolume ($\pi r^2 h$). Biomass was calculated as Carbon converting biovolume using (Menden-Deuer and Lessard 2000) formula; log_{10} C = -0.541 + 0.811 x log_{10} V. The converted log_{10}C gives pg of C $cell^{-1}$, which multiplied by maximum cell density ml^{-1} gives the maximum biomass attained in terms of pg of C ml^{-1}.

5.2.4 Oxylipin analysis. Culture pellets collected for oxylipin analysis were extracted and analysed according to the methods described in (d'Ippolito *et al.* 2009) and briefly explained in Chapter IV. Oxylipin analysis is presented only for the *L. danicus* strain since the *L. aporus* strain that had been selected for the whole transcriptome analysis and hence used in this study, on subsequent analysis, showed no oxylipin production.

5.3 Results

5.3.1 *Acclimation*

Acclimating pre-treatment of cultures to lower and higher T allow for the cells to gradually change their morphological, physical, and/or biochemical traits in response to changes in

their environment, possibly avoiding stress responses. Acclimation is plastic and reversible process during which diverse physiological modifications occur that usually terminate when the environmental condition stabilize. After the physiological and biochemical adjustments, the cells maintain a constant μ_{max} provided that other factors remain constant. In order to test whether the cells were acclimated to the given conditions, four sequential growth tests were assessed. The μ_{max} of the eight acclimated strains are represented in Table 5.2. The growth curves and maximum growth rates of cultures of all the four strains of both species were fluctuating in the initial two cycles of the growth test. In the subsequent two growth tests, the growth curves were more stable and μ_{max} were similar. When the slopes and μ_{max} of last two consecutive transfers were not significantly different the cultures were considered acclimated. Therefore the initial 35 days of acclimation to T conditions was sufficient for the cell to balance the cellular mechanisms towards the changes in T.

5.3.2 *Effect of temperature on growth and physiological characteristics*

Growth curves for *L. danicus* are presented in Fig. 5.1 (A–D) as cell number counts and in Fig. 5.2 (A–D) as fluorescence measurements, while for *L. aporus* they are presented in Fig. 5.3 (A–D) as cell numbers and in Fig. 5.4 (A–D) as fluorescence measurements. μ_{max} values calculated from cell multiplication rate during the exponential phase as a measure of physiological response to variation in T for four strains of *L. danicus* and *L. aporus* are presented in Fig. 5.5, A. The maximum cell density attained, the number of days taken to reach the stationary phase and the total biomass produced for each of the

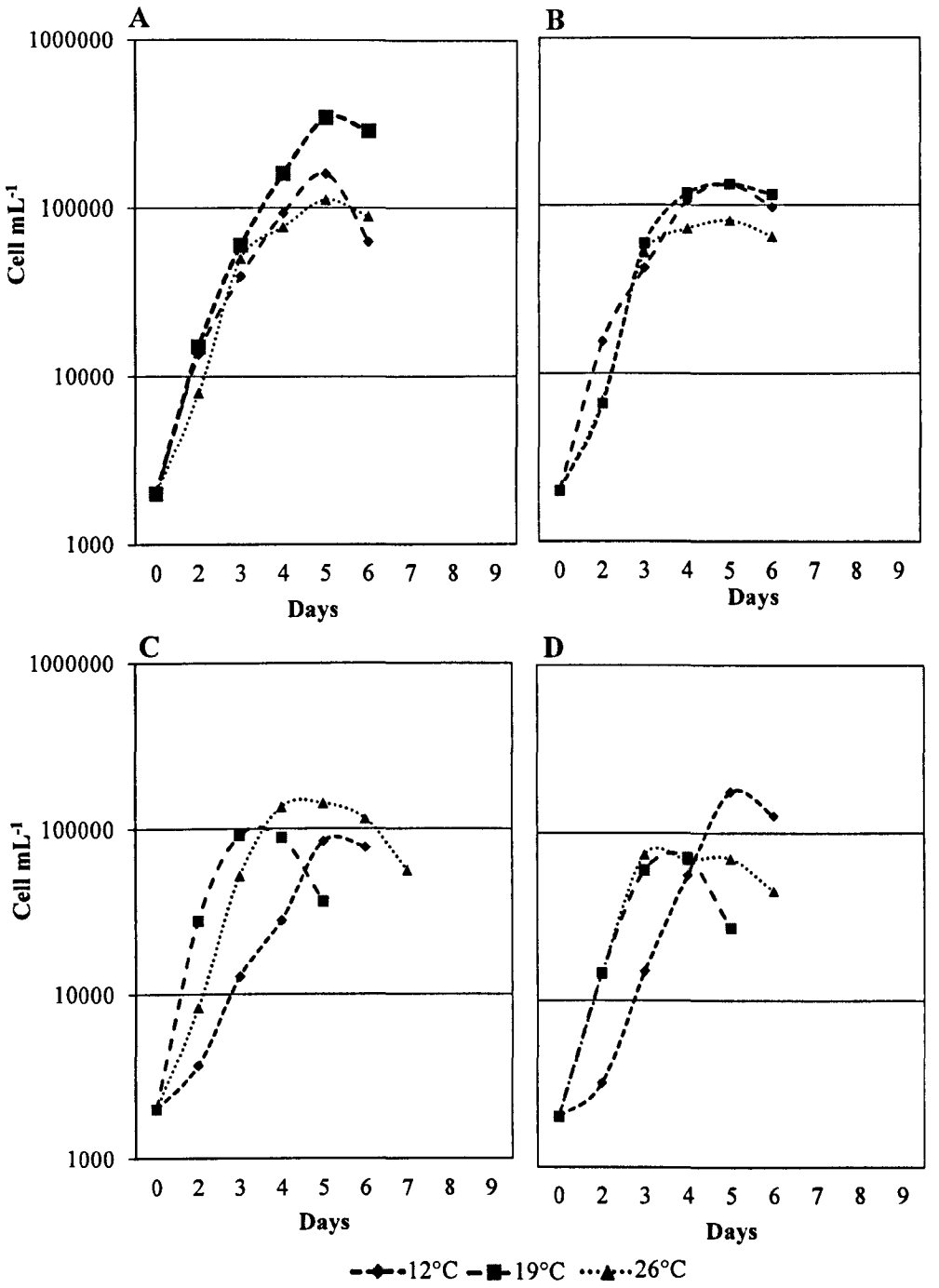


Fig. 5.1. *Leptocylindrus danicus* growth curves established by daily cell density counts at three different temperatures (n=1). (A) strain SZN-B707. (B) strain SZN-B714. (C) strain SZN-B715. (D) strain SZN-B650

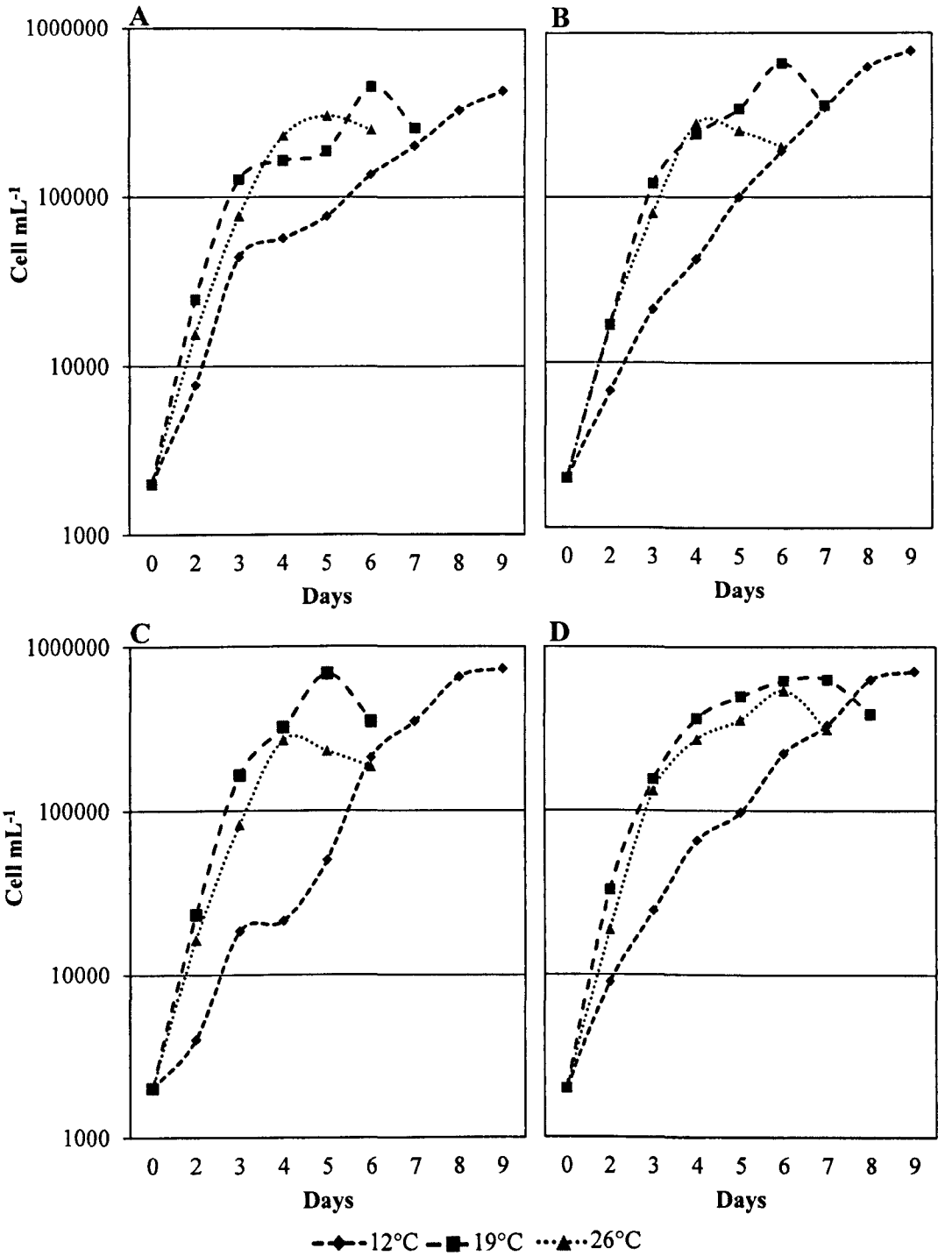


Fig 5.2. *Leptocylindrus aporus* growth curves established by daily cell density counts at three different temperatures (n=1). (A) strainSZN-B727. (B) strain SZN-B651. (C) strain SZN-B752. (D) strain SZN-B753

species strain at three growth T values are presented in Figures 5.5, B–D, respectively. Results for maximal growth rates and biomass production are calculated based on the cell number counts.

i. *Maximum growth rate.* Based on cell counts, At the T of 19 °C *L. danicus* grew exponentially for 3–4 days, followed by slow growth for a day or two, and finally a decline phase (Fig. 5.1, A–D). Likewise, at 19 °C, the species *L. aporus* grew exponentially for 4–5 days, followed by a day or two of slow growth, later entering a decreasing growth phase (Fig. 5.2, A–D). At 26 °C and 12 °C, deviation from the normal growth cycle was clearly visible but the growth curves at these Ts were always below the curve of 19 °C, at times they overlapped.

The cultures of *L. danicus* had a very similar μ_{max} at the three T values, among which the one at 26 °C showed a slightly higher average μ_{max} (Table 5.3), while *L. aporus* had similar growth curves at both the higher T values (19 and 26 °C; Fig. 5.2, Students t-test $p<0.002$), and in some strains slightly higher growth rates at 19 °C (Fig. 5.2, C and D). At 12 °C, *L. aporus* showed significantly lower growth but grew steadily for a long time (Fig. 5.2; $p<0.002$). In both the species, there was an increase in the μ_{max} with the increase in the T (Fig. 5.5, A) from 12 °C to 19 °C. The increase was far more evident in *L. aporus*, whereas *L. danicus* the difference was lower because the different strains had a relatively high μ_{max} (1.3 d^{-1}) even at 12 °C. At this T value, the differences in growth between the two species was the highest, whereas μ_{max} values were rather comparable between the two species, considering all strains, at 19 °C and 26 °C ($p>0.02$). Indeed, the increase in temperature from 19 °C to 26 °C did not have a remarkable effect on the μ_{max}, although at 26 °C all strains of *Leptocylindrus danicus* showed a lower growth rate (1.3 $d^{-1}\pm0.07$) than at 19 °C (1.6 $d^{-1}\pm0.15$; $p=0.01$) while in the case of *L. aporus*, two strains had a higher μ_{max}, one a comparable μ_{max} and finally one a lower μ_{max} than at 19 °C, resulting in an average growth

rate of 1.5 d^{-1} at both T values (Fig. 5.5, A; $p>0.5$). The same trend was repeated during the growth and metabolomic experiment (Fig. 5.6, A).

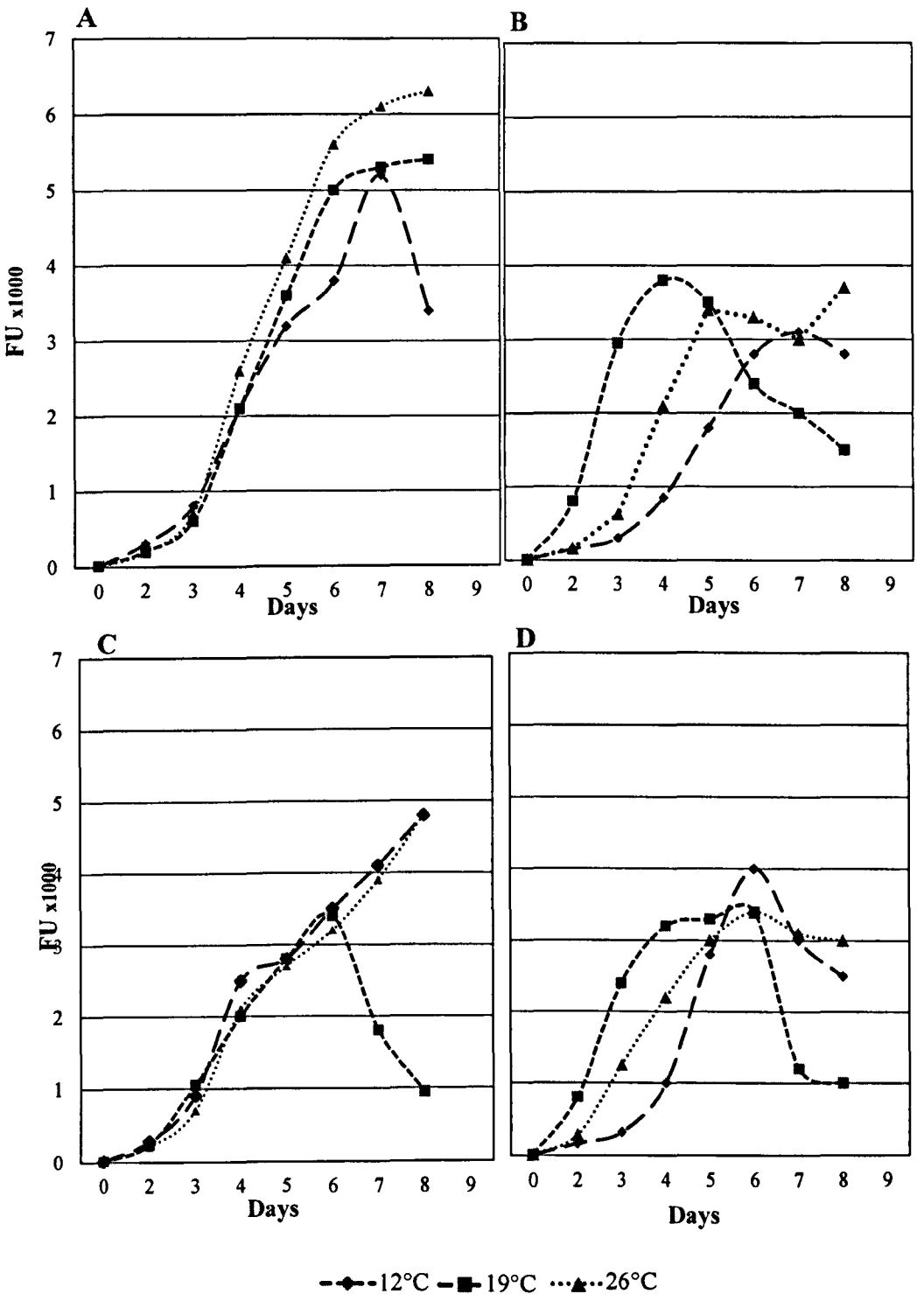


Fig. 5.3. *Leptocylindrus danicus* growth curves established by daily arbitrary fluorescence measurements at three different temperatures (n=1). (A) strain SZN-B707. (B) strain SZN-B714. (C) strain SZN-B715. (D) strain SZN-B650

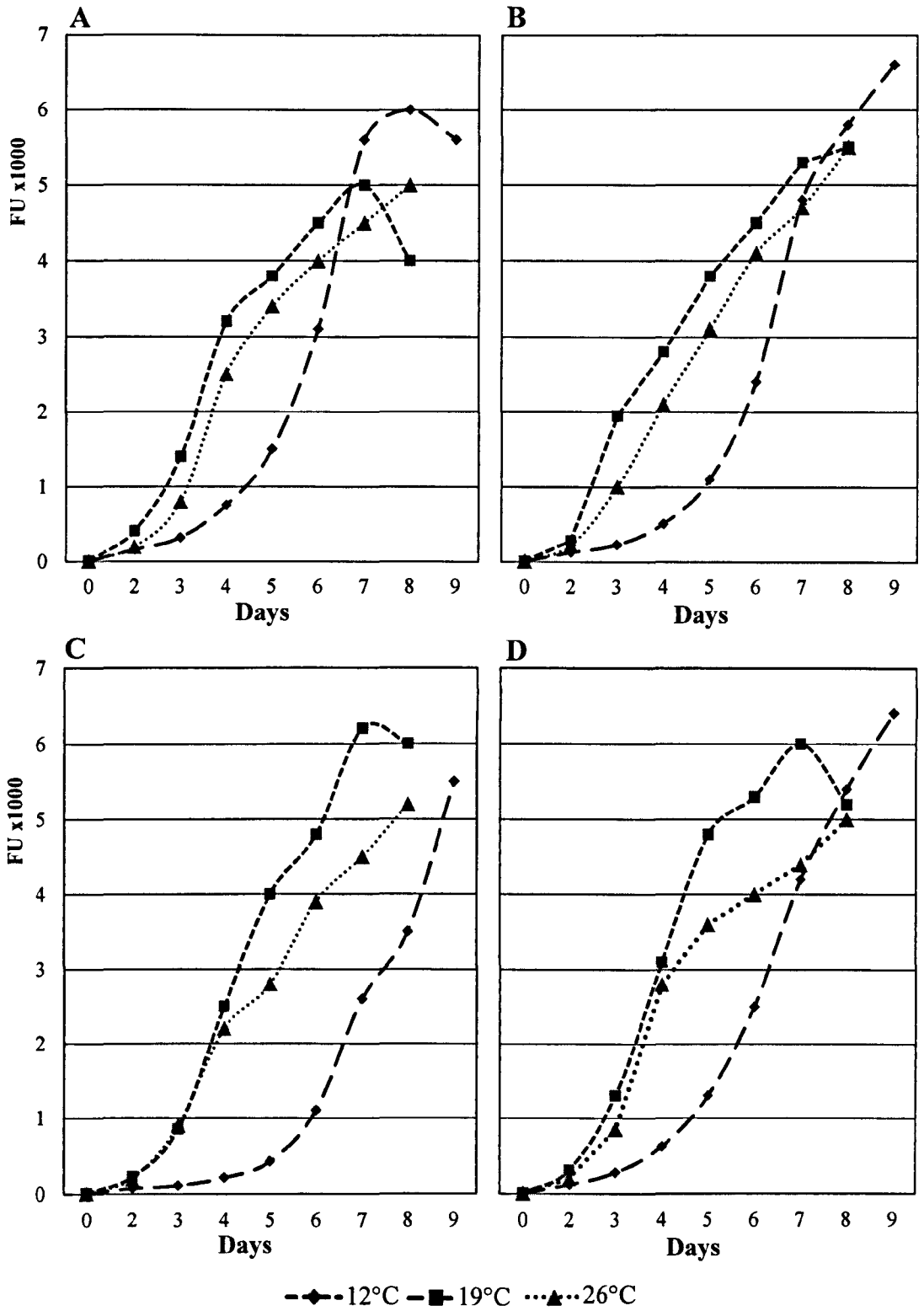


Fig. 5.4. *Leptocylindrus aporus* growth curves established by daily arbitrary fluroscence measurements at three different temperatures (n=1). (A) strain SZN-B727. (B) strain SZN-B651. (C) strain SZN-B752. (D) strain SZN-B753.

Table 5.2. Maximum growth rates (d^{-1}) for the four curves of the acclimation test inferred from cell counts (n=1). (A) 12°C. (B) 19°C. (C) 26°C.

A

	Leptocylindrus danicus				*Leptocylindrus aporus*			
12 °C	**B707**	**B714**	**B715**	**B650**	**B727**	**B651**	**B752**	**B753**
curve 1	1.09	1.29	1.27	0.86	1.21	0.96	1.34	1.15
curve 2	0.94	1.10	0.72	0.77	1.15	0.91	1.09	1.11
curve 3	1.35	1.54	1.47	1.74	1.19	1.15	1.22	1.19
curve 4	1.40	1.46	0.91	1.60	1.29	1.12	0.93	1.26
Avg	**1.23**	**1.37**	**1.03**	**1.37**	**1.21**	**1.06**	**1.08**	**1.19**

B

19 °C	**B707**	**B714**	**B715**	**B650**	**B727**	**B651**	**B752**	**B753**
curve 1	1.25	1.40	1.25	1.10	1.54	1.69	1.81	1.65
curve 2	1.46	1.31	1.17	0.84	1.26	1.31	1.30	1.46
curve 3	1.69	1.70	1.33	1.38	1.63	1.85	1.99	1.96
curve 4	1.60	1.54	1.46	1.37	1.68	1.80	1.91	1.93
Avg	**1.58**	**1.52**	**1.32**	**1.19**	**1.52**	**1.65**	**1.73**	**1.78**

D

26 °C	**B707**	**B714**	**B715**	**B650**	**B727**	**B651**	**B752**	**B753**
curve 1	0.85	0.64	0.75	0.73	1.52	1.56	1.54	1.61
curve 2	1.20	1.19	1.19	1.01	1.52	1.62	1.70	1.78
curve 3	1.43	1.60	1.65	1.49	1.77	1.78	1.81	1.92
curve 4	1.39	1.39	1.56	1.39	1.73	1.79	1.79	1.85
Avg	**1.34**	**1.39**	**1.47**	**1.29**	**1.67**	**1.73**	**1.77**	**1.85**

Based on an arbitrary fluorescence unit scale (FU), at 19 °C FU values in *L. danicus* increased on an average for 6 days (4–7 days), followed by a day of slow or sharp decline in the values (Fig. 5.3, A–D). Likewise, at 19 °C, FU values for the species, *L. aporus* increased exponentially for 7 days, followed by a day or two of marginal increase, later declining abruptly (Fig. 5.4, A–D). At 26 °C and 12 °C, a deviation from the normal growth cycle was clearly visible but the growth curves at these Ts were always below the curve of 19 °C, at times overlapping. The only exception was *L. danicus* SZN-B707, whose growth curve at 26 °C was higher than those 12 °C and 19 °C.

Table 5.3. Physiological growth responses expressed as (A) growth rates (d^{-1}) (B) maximum cell density (cells ml^{-1}) and (C) days needed to reach stationary phase to different temperature of acclimated cultures (n=1).

A

	Leptocylindrus danicus				*Leptocylindrus aporus*			
K	**B707**	**B714**	**B715**	**B650**	**B727**	**B651**	**B752**	**B753**
12 °C	1.30	1.46	1.07	1.35	0.78	0.93	0.93	0.90
19 °C	1.52	1.54	1.84	1.61	1.25	1.42	1.74	1.41
26 °C	1.24	1.39	1.36	1.39	1.54	1.52	1.50	1.39

B

Cell Den.	**B707**	**B714**	**B715**	**B650**	**B727**	**B651**	**B752**	**B753**
12 °C	159000	116000	84000	175000	428000	774000	725000	702000
19 °C	342000	133000	91000	71000	450000	650000	683000	627000
26 °C	111000	81000	143000	70000	301000	280000	270000	540000

D

Days	**B707**	**B714**	**B715**	**B650**	**B727**	**B651**	**B752**	**B753**
12 °C	5	5	5	5	10	10	10	10
19 °C	5	5	4	4	6	6	5	7
26 °C	5	5	5	5	5	5	5	6

ii. ***Cell density.*** The maximum cell density (C_{max}) attained is presented in Fig. 5.5, B for each strain of the two species. Strains of *L. danicus* attained a C_{max} of 7–34 $x10^4$ cells ml^{-1}, which was comparatively lower than in *L. aporus*, which attained a C_{max} of 27–77 $x10^4$ cells ml^{-1} ($p<0.001$ at 12 and 19 °C and $p<0.01$ at 26 °C). Although there is a small overlap in the range, the C_{max} attained at each of the three T values was significantly lower for *L. danicus* compared to *L. aporus*. Strains of *L. danicus* showed similar C_{max} at the three T values, with a tendency to increase at 19 °C compared to 12 °C and 26 °C, and with a higher standard deviation at this latter T. Instead, *L. aporus* showed a negative relationship with T, i.e. C_{max} decreased with the increase in T. The decrease was more pronounced at 26 °C, with larger overlaps at 12 and 20 °C among the strains. One strain showed a marginal increase, another showed a moderate decrease and two others showed a marginal decrease in C_{max} at 19°C compared to 12°C. The same trend was repeated during the growth and

metabolomic experiment (Fig. 5.6, A) in both the species. In *L. aporus*, the difference in C_{max} among the three T conditions overlapped in replicates ($p>0.02$).

iii. ***Days to complete the exponential growth phase.*** The total number of days taken for the completion of the exponential and attain the stationary phase is presented in Fig. 5.5, C. *Leptocylindrus danicus* took approximately 5 days to complete the exponential growth phase at all the three T ($p>0.1$), with an exception of two strains taking 4 days at 19°C. Contrastingly, *L. aporus* showed a decreasing number of days to reach the stationary phase with the increase in T. The difference was remarkable between 12°C and 19°C ($p<0.001$) whereas often the species took an extra day to complete the growth cycle at 19°C compared to 26 °C. The only exception was the strain SZN-B752, which took 5 days at both T conditions.

During the growth and metabolomic experiment with a single strain of each species, the two species showed the same trend as in the multiple strain experiment, but they took one day less to complete the growth cycle (Fig. 5.5, C and Fig 5.6, C). Additionally, *L. aporus* took 6 days instead of 10 days to complete the growth cycle. For *L. danicus,* the regression analysis shows that for each one-degree rise in temperature μ_{max} increases by 0.056 d^{-1}, C_{max} decreases by 126 cells ml^{-1} and the time to reach stationary phase increases by 0.07 days. On the other hand, for the same T variation, *L. aporus* μ_{max} increases by 0.08, C_{max} decreases by 4800 cells ml^{-1} and the time to reach stationary phase increase by 0.2 days. The two species also showed a negative regression coefficient for biomass (0.007 and 0.22) and a positive regression coefficient for carbon content per cell (1.5 and 0.32).

iv. ***Cell size and morphology.*** During the experiment, *L. danicus* had a diameter of ca 8 µm and *L. aporus* had a cell size of ca 5 µm during exponential phase, without any noticeable changes among the three temperatures. In the growth curves built for the metabolomic experiment more detailed observations were made for both the species, and a variation in the pervalvar cell length with temperature was observed (Fig. 5.6, D). *Leptocylindrus*

danicus showed an increase in average pervalvar distance with the increase in T and this was evident between 12°C or 19°C and 26°C but not between 12°C and 19°C (Fig. 5.6, D). The average pervalvar distance in *L. aporus* was the highest at 19 °C, while it decreased at 12 °C and 26 °C (Fig. 5.6, D). Yet, in both the species the range in the pervalvar distance of cell in exponential phase overlapped, with no clear trend.

During the experiments, no significant changes in the morphology of the two species were observed in either species. In both species, cells at 12 °C tend to form longer colonies than at higher Ts. This change was more pronounced in *L. danicus* strains, while *L. aporus* showed this tendency in a less intense way. However, these results were based on long-term visual observations and no precise quantification was made. Another long-term observation was that the species *L. danicus* needed frequent and precise intervals of subcultures at 26 °C whilst *L. aporus* did not withstand long-term culturing at 12 °C. This indicates that *L.* aporus might experience physiological stress at the lowest T conditions.

V. Biomass. The maximum biomass build up for the four strains of each *L. danicus* and *L. aporus* is presented in the Fig. 5.5, D and Fig. 5.6, E. The species *L. aporus* produced approximately twice the biomass than that of *L. danicus* at all the three Ts. In both species, the strains did not show a clear trend, except for one strain of *L. danicus* and two strains of *L. aporus*, which showed a decrease in biomass with increase in T. On an average among the strains, *L. danicus* produced 18.7, 15.3 and 11.2 µg of C ml^{-1} and *L. aporus* produce 37.8, 40.3 and 27.5 µg of C ml^{-1} at 12, 19 and 26 °C, respectively. These values show that in *L. danicus* there was a decrease in biomass accumulated with the increase in T, while in *L. aporus* there was a slight increase in biomass accumulated from 12 °C to 19 °C and a remarkable decrease at 26 °C compared to the other two T conditions. In *L. aporus,* during the growth and metabolomic experiment with a single strain and five replicates, similar results were observed to that of the average among the strains, whereas in *L. danicus* the

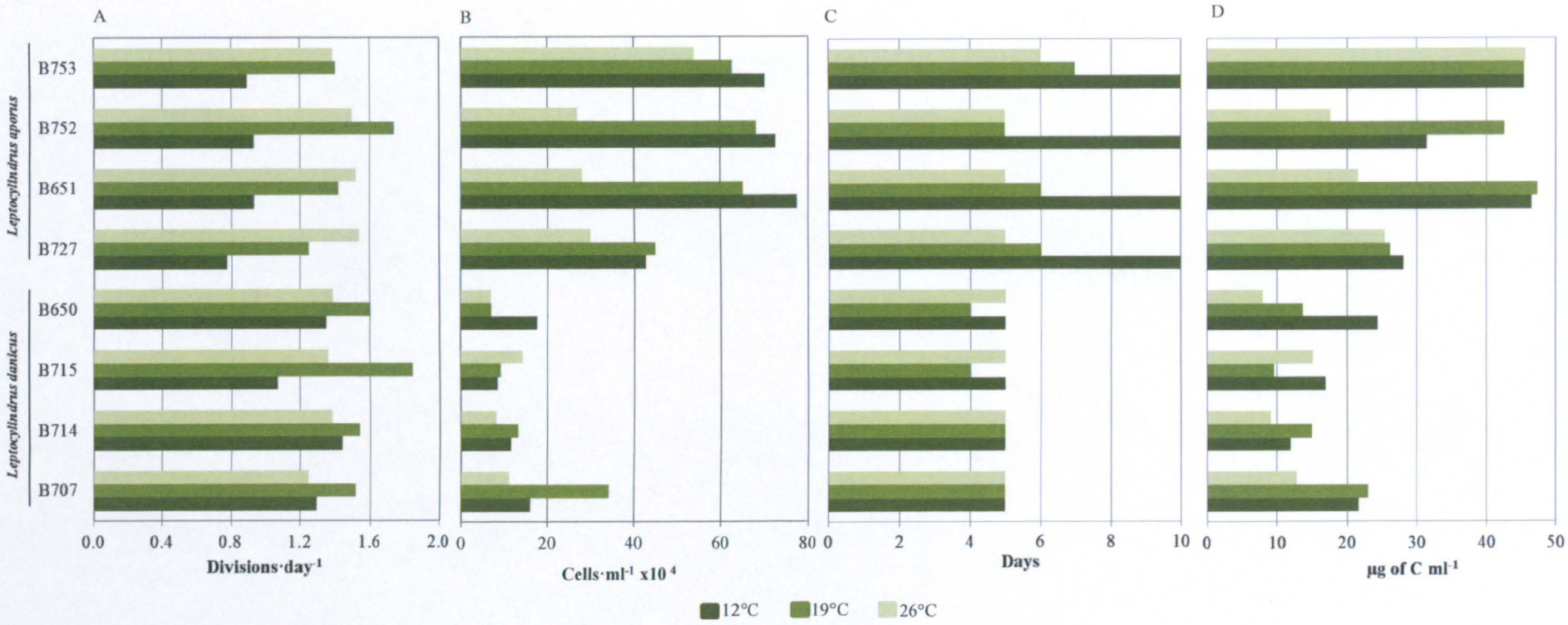


Fig 5.5. Physiological responses of *L. danicus* strains and *L. aporus* strains to three different growth temperatures (n=1). (A) Maximum growth rates; calculated based on cell counts. (B) Maximum cell density attained at the end of the experiment. (C) Number of days taken to complete growth cycle. (D) Biomass produced at the end of the growth cycle.

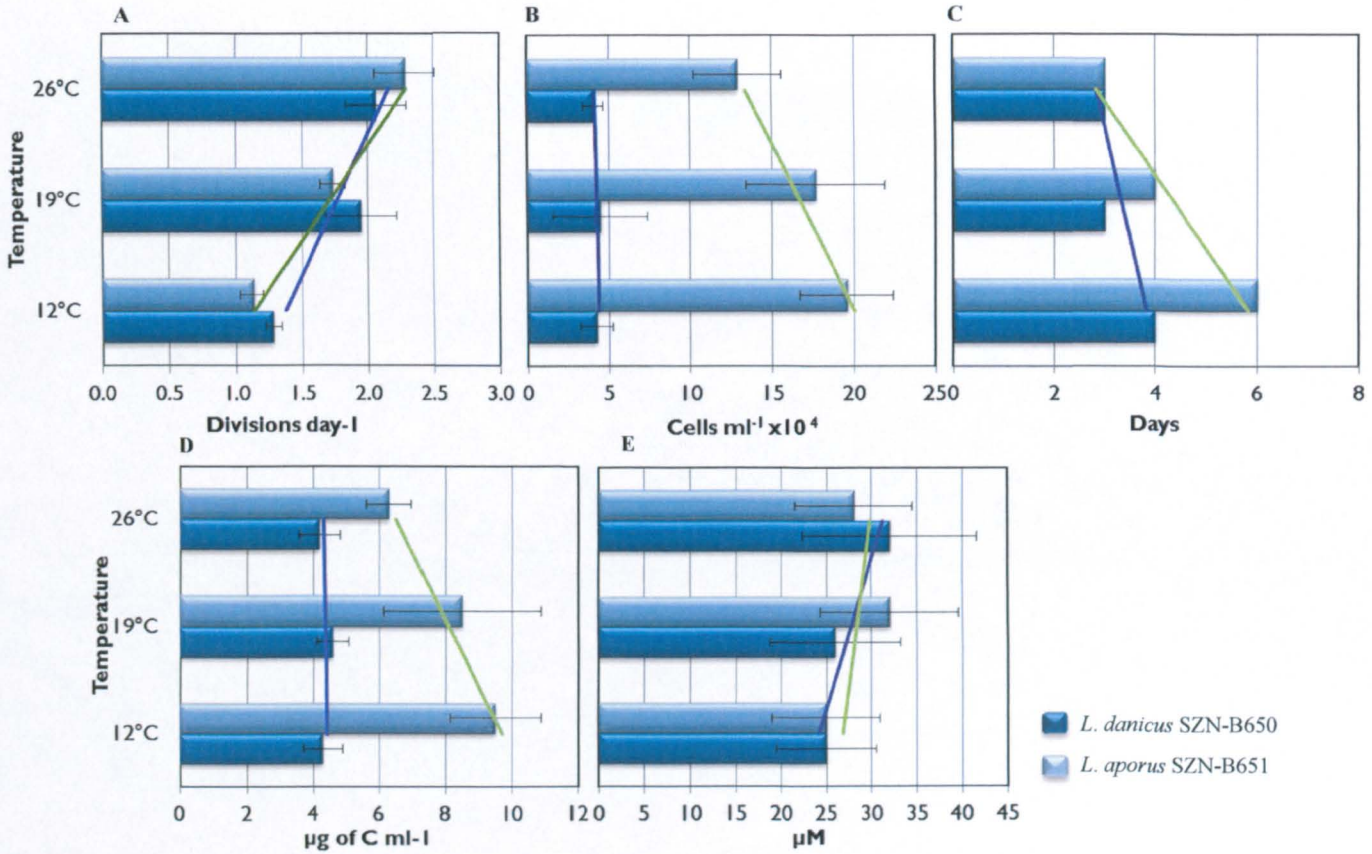


Fig. 5.6. Physiological responses of *L. danicus* and *L. aporus* to three different growth temperatures (n=5). (A) Maximum growth rates. (B) Maximum cell density attained at the end of the growth curve. (C) Number of days taken to complete growth cycle. (D) Biomass produced at the end of the growth cycle. (E) Pervalvar distance. Linear trend curves are included to approximate the responses.

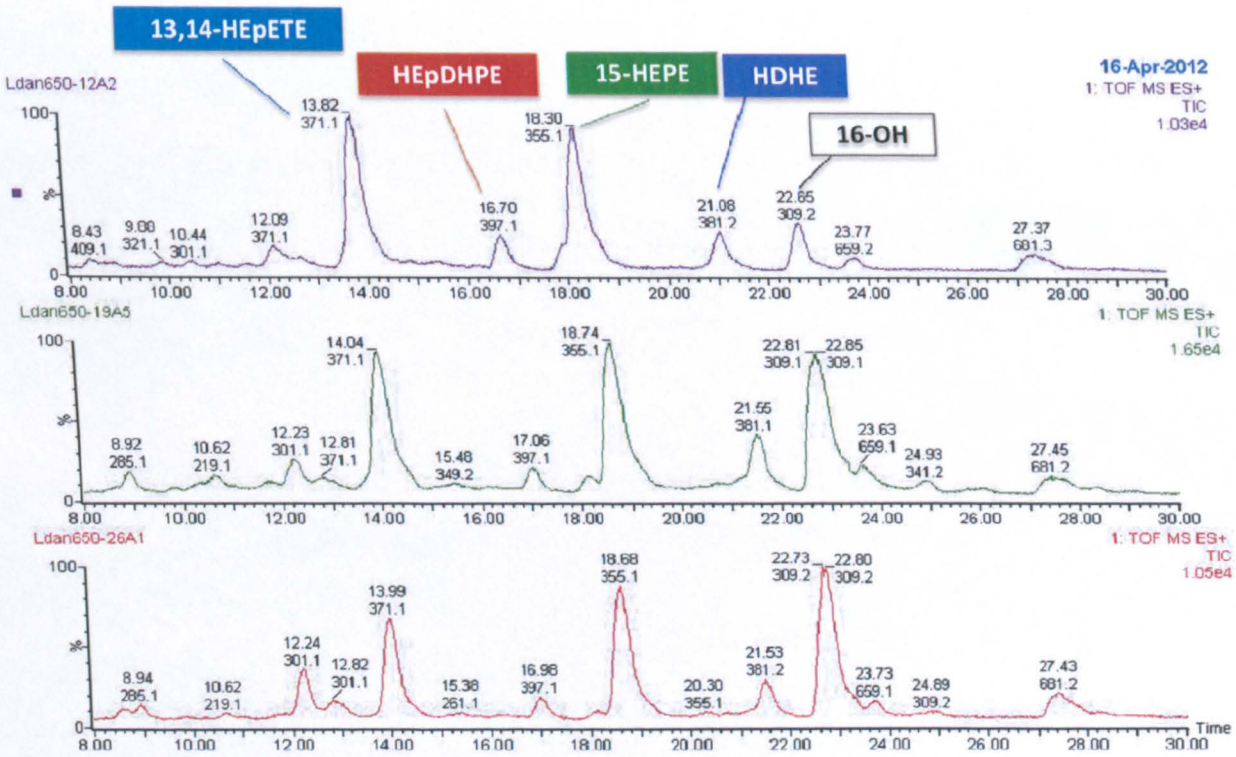
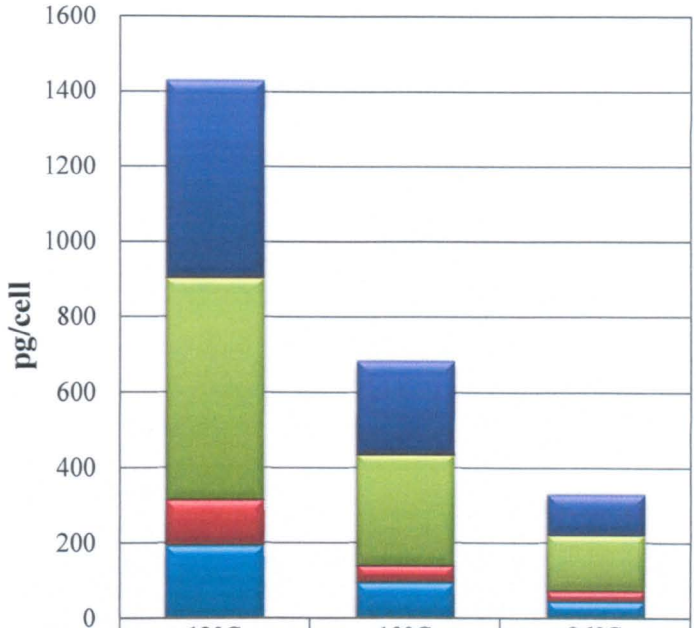


Fig. 5.7. Typical oxylipins profiles of *L. danicus* SZN-B650 exposed to three different temperature conditions.



	12°C	19°C	26°C
13,14-HEpETE	525.79	246.16	108.52
15-HEPE	589.99	296.45	149.84
HEpDHPE	120.64	46.39	26.77
HDHE	194.88	95.36	46.42

Fig. 5.8. LC-MS quantifications of oxylipins produced by *L. danicus* SZN-B650 exposed to three different temperature conditions

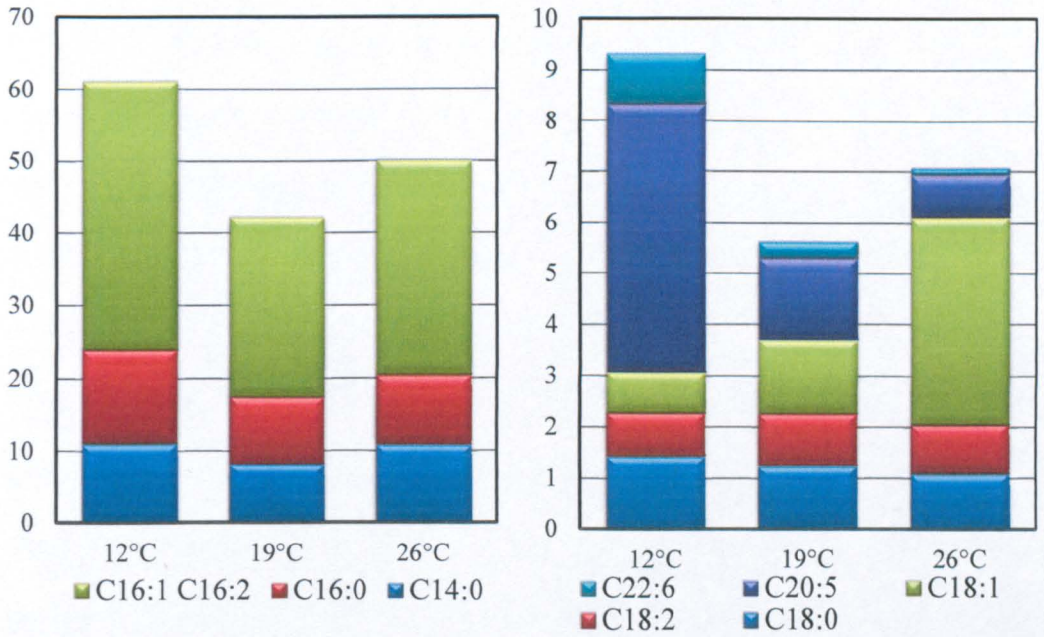


Fig. 5.9. GC-MS quantifications of fatty acids produced by *L. danicus* SZN-B650, under three different temperature conditions.

species produced a similar biomass at the end of the stationary phase at all the three Ts. Nevertheless, between the two species, the biomass build-up in *L. aporus* was twice of L. *danicus* except at 26 °C.

5.3.3 *Oxylipin and fatty acid profiles*

Leptocylindrus aporus SZN-B651 did not produce oxylipins in any of the exposed temperature. The typical LC-MS chromatograms for oxylipins produced by the species *L. danicus* at the three T conditions are presented in Fig. 5.6. Two pathways of oxylipins synthesis were recognised in the species (Fig. 5.7, Chapter IV). These involve the oxidation of C_{20} and C_{22} fatty acids to hydroxide reductase derivatives 15-HEPE and HDHE respectively and epoxyalcohol synthase derivatives 13,14-HEpETE and HEpDHPE respectively (Fig. 5.8, A and B). All the four compounds were detected at the three T condition. However, the quantity of oxylipins produced decreased with the increase in T (Fig. 5.8). Nevertheless, the percent levels of different oxylipins were comparable, i.e. the proportionality of the four oxylipins remained similar across the three Ts (Fig. 5.8). To comprehend the reason for the higher amount of oxylipins production at lower growth T, the fatty acids, which are the precursor molecules of oxylipins, were quantified. GC-MS analysis of fatty acid content of the samples showed that *L. danicus* produces C16, C18 and C20 saturated fatty acids (SFA) and C16, C18, C20, and C22 unsaturated fatty acids (USFA). The species produced 25.5 pg $cell^{-1}$ and 45.05 pg $cell^{-1}$ of SFA and USFA at 12°C. At 19°C and 26°C the species produced 18.73 pg $cell^{-1}$ and 29.12 pg $cell^{-1}$ and 21.65 pg $cell^{-1}$ and 35.6 pg $cell^{-1}$ of SFA and USFA, respectively. The variation in the SFA content was not profoundly different at three Ts whereas the USFA produced at 12°C was significantly higher than at 19 °C and 26 °C. Together the SFA and USFA specifically the precursor fatty acids C20:5 and C22:6, produced decreased with the increase in T (Fig. 5.9 A and B). Taken together, *L. danicus* produce a higher quantity of fatty acids (C20:5 and C22:6) and oxylipins at 12 °C and both the compounds decrease with increase in T.

5.4 Discussion

The aim of this study was to comprehend the phylogenetic and ecological background influence on the physiological responses of two *Leptocylindrus* species at different temperatures. *Leptocylindrus danicus* and *L. aporus* are the member species of the two genetically distinct lineages within the genus *Leptocylindrus* (Chapter II). The two species strains were isolated from the same geographical area but were found to bloom under contrasting environmental conditions (Chapter III). The two species were also found to have opposite life cycle patterns with *L. danicus* having sexual reproduction and auxospore formation to restore cell size while *L. aporus* shows cell size restoration by vegetative auto-enlargement (French III and Hargraves 1986, Chapter II). Both the species were cultured under a wide range of temperature (12 °C, 19 °C and 26 °C) to reveal differences in their physiology and metabolite composition.

On exposure to the three tested Ts, the species *L. danicus*, with slightly lower μ_{max} at 12 °C and similar μ_{max} at higher Ts, similar C_{max} at all Ts and an extra day at 12 °C with respect to higher Ts to reach the stationary phase, can produce a similar biomass at the end of the experiment. By contrast, *L. aporus* showed variations in the measured parameters leading to variations in the yield biomass at the end of the growth cycle. At 12 °C, the species, with lower μ_{max}, in a longer time period could attain the highest C_{max}, producing higher biomass. At 19°C, with higher μ_{max}, much less time to complete the growth cycle than at 12 °C, could attain a higher C_{max}, producing the highest biomass. At 26 °C, with the highest μ_{max}, far lower growth period than at 12 °C, attained the lowest C_{max}, producing the lowest biomass. Thus, it appears that *L. danicus* cellular mechanisms are adapted to grow

under wider Ts than *L. aporus*, which is equally efficient at higher temperature but less at low temperature conditions.

Physiological responses

5.4.1 ***Maximum growth rates as a function of temperature.*** The overall effect of the biotic and abiotic factors on cell physiology is reflected in the growth rate and finally in the biomass accumulation. These parameters have been used for several decades as metrics to understand the effect of environment, based on *in vitro* experiments. Growth rate variation as a function of temperature is hypothesized to be either exponential, with growth rate doubling with every 10 °C increase (following an Arrhenius relationship, Q_{10}=2, where Q_{10} is the factor by which a biological rate is increased following a 10 °C rise in the temperature; van't Hoff 1884, Eppley 1972, Raven 1988), or showing a linear slope. Further, bell shaped responses have also been reported (Fiala and Oriol 1990). The results here presented in the study show a linear and exponential increase in the maximum growth rate with increase in the temperature for the two species examined. The species, *L. danicus* showed a linear increase in the growth rate from 12 °C to 19 °C (based on cell counts); while a further increase did not significantly increase the growth rate indicating cells may have suffered a negative effect of temperature increase. *Leptocylindrus aporus* showed an exponential growth rate trend with a marked increase in growth from 12 °C to 19 °C, but a further increase to 26 °C only showed a slight increase. Overall both species showed a Q_{10} of 2 for a temperature increase from 12 °C to 26 °C, demonstrating an exponential curve. In protists, growth rate is a combination of many processes and its increase with temperature is generally considered to be linear rather than exponential or power response , especially when single species are considered as compared to when multiple species are grouped in community studies (Montagnes 1996). Thus, maximum growth rate change in response to different controlling factors can be considered a species-specific character.

Interspecific variation in growth rate can be related to cell size with smaller diatoms generally having higher growth rates (Geider *et al.* 1986, Sarthou *et al.* 2005) because of the catalytic advantage. Although in the current study there were no changes in cell valve diameter, the changes observed in the pervalvar distance provide some indications on the physiological status of the cells. Under stress, cells tend to divide slower and hence have a larger pervalvar distance than cells growing under optimal conditions. The two species had different average pervalvar distance, although range overlapped. The increase in average pervalvar distance with increasing temperature in *L. danicus* indicates, the species prefers lower temperature for normal cell division. In contrast, the results for *L. aporus* are less clear, as the lower average pervalvar distance at 19 °C and higher at 12 °C and 26 °C indicate the species divides easier at 19°C. An increase in cell size decreases the surface area to volume ratio and subsequently the number of membrane transporters, thus reducing the solute influx and efflux. Besides, the chlorophyll-a specific absorption cross-section coefficient increase with cell size thus decreasing the effectiveness of photon absorption per chlorophyll molecule in larger cell than in smaller cell, the so-called 'package effect'. In simple terms the 'package effect' implies that the increase in cell size decreases the effectiveness of increased pigmentation in harvesting light (Kirk 1994, Finkel and Irwin 2000).

Growth rate calculations based on cell numbers are preferred over fluorescence based measurements as there might be the problem of self-shading of cell when they reach higher abundances, which can cause an increase in chlorophyll independent from cell divisions. Additionally, fluorescence yield also depends on the physiological status of the cell, which is more relevant during the stationary and decline phase of the growth curve as cells experience various stresses at this stages.

5.4.2 ***Cell volume and morphology as a function of temperature.*** The relationship between diatom cell size and temperature has been well debated in literature. Many studies have demonstrated a decreasing cell size with the increasing temperature (Jørgensen 1968, Margalef 1989, Montagnes and Franklin 2001) and still others that do not provide a clear trend (Yoder 1979, Sournia 1982). There are also reports of cell size reduction to decrease in Fe availability (Kudo *et al.* 2000); hence it is important to note this change in the study of response to niche. In the current study both the species visually appear bigger cells (but not in measurements) in their preferred temperatures. This means, *L. danicus* which grows well at lower temperature also visually appear vigorous at those temperature values. On the contrary, *L. aporus*, which prefers warmer temperature, visually appear vigorous at 19 °C, and at 26 °C, cells appeared to be in better status than those of *L. danicus*. However, no remarkable and clear trend for change in cell size and volume was observed. These observations are in line with Verity (1981) who demonstrated no clear trend for *L. danicus*, although it is difficult to trace which of the species described in Chapter II were used in their experiments. The possible explanations to this kind of slight variation or no clear trend include the interplay of the growth rate and cell size expansion (Montagnes and Franklin 2001). These two components control when and how changes in cell size occur. A higher growth rate at higher temperatures might decrease the average cell size of cultures which may need to undergo cell size expansion earlier as compared to the slow growing cultures. Another factor influencing cell size is life cycle pattern of a species i.e. whether the species undergoes sexual or vegetative process to restore the cell size. In *L. danicus* and *L. aporus*, the sexual reproduction in *L. danicus* and vegetative reproduction in *L. aporus* has been demonstrated to be clearly dependent on temperature (French III and Hargraves 1985, French III and Hargraves 1986). Evidently, further careful investigations are needed to understand whether diatom cell volume can be influenced by temperature

and this is a relevant issue as cell size can affect sedimentation rate and autotrophic production.

The colony size i.e. the number of cells per colony decreased for both the species with the increase in the temperature and this was growth stage specific, with the number of cells per colony being higher in early stages (lag and early exponential phase), and hence decreasing towards the stationary phase. Seasonal changes in colony size have been noticed in natural samples (Lund *et al.* 1963) and have also been demonstrated in laboratory studies (Hayakawa *et al.* 1994, Werner and Stangier 1976). There is also confusion in understanding the trend in this feature in response to environmental changes as there are studies that demonstrate a decrease in colony size with increase in temperature (Werner and Stangier 1976) and other that demonstrate the contrary (Hayakawa *et al.* 1994). There are still other reports of combined effects of different niche factors, including changes in nutrient concentrations, for example silica and temperature (Werner and Stangier 1976). Colony formation has important ecological implications as it determines the sedimentation rate and grazing pressure from zooplankton.

5.4.3 ***Cell number and biomass as a function of temperature.*** Biomass production is directly dependent on the physiological status of the cell and is thus influenced by abiotic factors. In the study, for *L. aporus* we observed approximately double biomass accumulation as compared to that of *L. danicus*. It can be a species-specific character that certain species can reach higher cell abundance than others. The decreasing biomass demonstrated by *L. aporus* was clearly the response to temperature which is the result of decreasing cell number with the increase in the temperature.

The variation in growth rate, cell density and subsequent biomass accumulation is dependent primarily on the photosynthetic ability of cells. As photosynthetic carbon assimilation is enzymatically controlled, it is a temperature-dependent process. It is

frequently assumed that the light-dependent portion of the photosynthetic response is temperature independent but this is proposed not to be strictly true (Raven and Geider 1988) for two reasons. Firstly, electron transfer process is dependent on membrane fluidity and, secondly, the electrochemical turnover of plastoquinone and plastocyanin do depend on temperature (Falkowski and Raven 1997). At low temperature, even at relatively low irradiance levels, the limitation of electron transport or carbon fixation reduces the cell's ability to absorb light. The resulting excess energy may cause photoinhibition due to damage to the PSII apparatus. In general many aquatic photoautotrophs acclimate to changes in temperature in a comparable fashion to that of photoacclimation. Temperature dependent activity and level of expression (Bose *et al.* 1999) of Ribulose-1,5-bisphosphate carboxylase/ oxygenase (RuBisCo), the key enzyme involved in carbon assimilation in photosynthetic organisms, thus regulating the ability of cells to produce and store carbohydrates. In *Chlamydomonas reinhardtii,* variation in RuBisCo content has been reported as an adaptive mechanism at low temperatures to counterbalance poor catalytic efficiency (Devos *et al.* 1998). Under high temperature, substrate affinity of RuBisCo increases more slowly for O_2 than for CO_2 with the consequence that the potential for photorespiration increases (Berry and Raison 1981); this is amplified by temperature dependent changes in the relative solubility of CO_2 and O_2 (Raven and Geider 1988). Together, increased concentrations of PUFAs in lipids of the thylakoid membrane to maintain electron transport under low temperatures was shown in the psychrophilic *Chlamydomonas subcaudata* (Morgan-Kiss *et al.* 2002). These alterations lead to a reduction in the light absorption capacity while increasing photosynthetic capacity. The effect is that light-saturated photosynthetic rates per unit carbon biomass can often be maintained at decreased temperatures, while simultaneously the propensity of the cell or organism to photoinhibition is also maintained (Falkowski and Raven 1997). The photosynthetic process is also closely associated with lipid biosynthesis, since mobilization

of lipids is partly controlled by carbohydrate levels as demonstrated by Larson and Rees (1996) in *Phaeodactylum tricornutum*. It is also viewed that at high temperature cell shift from light saturated condition to light limited condition, thereby experiencing a light limitation stress rather than a temperature stress. This might explain the negative relationship of *L. aporus* cell density with T. Hence, temperature determines major biological activities and finally biomass accumulation. Cells may mitigate the negative effects of lower temperatures through many physiological processes by either making more of an enzyme or by using an alternative form of the enzyme with higher kinetic efficiency (Raven and Geider 1988). This might explain the marginal differences in the observed values of the μ_{max}'s of *L. danicus* at the three Ts.

Biochemical Changes

Biochemical composition of microalgae including lipids, protein and carbohydrates content can be strongly influenced by temperature by effecting cellar processes. There alterations may reflect an adaptive response to perturbations in the environment. Responses to environmental stresses occur at all levels of cellular organization including adjustment of the membrane permeability, modifications of the cell wall architecture, and changes in cell cycle and cell division, additionally in production of compatible solutes (e.g. proline, raffinose, and glycine betaine) and restoration of cellular redox balance. Taking into account the importance of diatoms as primary producers in the marine food web many studies have also concentrated on the changes in the nutritional properties with the variation in temperature.

5.4.4 ***Lipids, Polyunsaturated fatty acids as a function of temperature.*** Lipids and fatty acids play important structural and functional roles in cell membranes and cellular processes including photosynthesis, secretion, signal transduction, vesicle trafficking, cytoskeletal recognition, and environmental responses. The cellular lipid composition in

microalgae is often affected by environmental factors, including temperature (Rousch *et al.* 2003, Renaud *et al.* 1995), salinity (Huflejt *et al.* 1990), light intensity (Solovchenko *et al.* 2008), nitrogen, phosphorus and silicon (Fidalgo *et al.* 1998 {Khozin-Goldberg, 2006 #13638) (Roessler 1988) (Yu *et al.* 2009). In general lipid content increase with the increase in culture temperature until a maximum value and later decrease (Thompson *et al.* 1992). In the current study we noticed significantly higher lipid content at lower temperature (12°C) and slightly higher content at higher temperature, indicating an inverted U shaped relationship with temperature, which are in concur with previous studies (Thompson *et al.* 1992, Sayegh and Montagnes 2011). This indicates that lipids are important at both higher and lower temperatures.

The decrease in the PUFA contents per cell with increase in temperature, specifically the highly unsaturated fatty acids (HUFA), C20:5 and C20:6 observed in this study is a common trend reported for microalgae. For example, the percentages of 20:5 in six marine algae, including *Chaetoceros*, *Rhodomonas*, *Isochrysis* and *Cryptomonas* were reported to be higher at lower temperatures (Renaud et al., 2002). Other reports of significant decrease in the production of HUFA at higher growth temperatures exist for species like *Chaetoceros calcitrans* (Thompson *et al.* 1992), *C. simplex* (Thompson *et al.* 1992) and *Nitzschia spp* (Renaud *et al.* 1995). However there are other reports where higher PUFA content at higher temperature, for e.g. *Stichococcus* (Teoh *et al.* 2004) and still other do not show a very clear trend in response to temperature (Renaud *et al.* 2002). Hence, it appears that the relationship between PUFA production and temperature is a species specific character (Renaud *et al.* 1995, Teoh *et al.* 2004).

Depending on the length and degree of unsaturation of fatty acid chains, temperature can have a profound effect on membrane fluidity (Harwood 1988). We did not observe this trend for *L. danicus*, where the proportionality of the SFA and USFA was similar at all the three temperatures. In many other species, a high growth temperature has been shown to

increase the saturated fatty acids (Thompson et al., 1992; Renaud et al., 1995) thus increasing the melting temperature (decreasing the fluidity) of the membrane. This suggests that the total lipid variation observed at three temperatures was sufficient to maintain the membrane fluidity of cells.

5.4.5 ***Oxylipins as a function of temperature.*** *Leptocylindrus aporus* SZN-B651 did not produce oxylipins hence the changes are noted only for *L. danicus*. This type of observations in strain difference is also made in other species. In fact other strains of *L. aporus* show that the species does produce oxylipins. The reason as to why some strains in a species don not produce oxylipins is rather unclear. Together fatty acids profile of *L. aporus* was not analysed for the strain because of lack of oxylipins in the strain. Moreover it is well established mechanism that lipid content varies from temperature that regulates membrane permeability. The observed inverse relationship of oxylipins to temperature *L. danicus* has also been reported for *Nanochloropsis* sp., where eicosapentaenoic acid content decreased with the increase in the temperature (Hu and Gao 2006). Although the observed variation in the total as well as in the individual oxylipin content among the three temperatures is in accordance with the variation in C20:5 fatty acid, it is difficult to attribute the variation to the non-availability of precursor molecules. Oxylipin biosynthesis is often regulated by environmental stimuli and developmental cues; their generalisation is indisputably difficult. Variation in oxylipins content is often attributed to the regulation at the release of fatty acids by phospholipase and the primary enzyme lipoxygenase involved in the first step of oxidation process. In addition, variation in the content across the growth cycle has been reported in *Pseudo-nitzschia delicatissima*, which shows increase in oxylipin content with the growth curve and growth stage specific compound, 15-oxoacid, produced only during stationary phase. (d'Ippolito *et al.* 2009). Nevertheless, in the current study the relationship between oxylipins and their variation and their adaptive significance is highly uncertain. Hence, further biochemical studies are

necessary to investigate the ecological significance of variation in the quantity but not the diversity of compounds produced in response to temperature stress.

5.4.6 ***Intraspecific variations in physiological and biochemical responses.*** Studies of physiological and biochemical parameters on multiple strains have clearly demonstrated the intraspecific variation in physiological responses (Balzano *et al.* 2011) and biochemical mechanisms (Gerecht *et al.* 2011). In the present study, marginal differences in physiological responses among the strains in both the species were noted. Differences were also observed when multiple replicates of same strain were studied indicating that all the strains had a range within which they responded. Strains taken from distinct environments might also show intraspecific variations in physiological responses (Balzano *et al.* 2011). In addition continuous laboratory cultivation under optimal cultivation, especially in species with short generation time, might allow for a selection process that generates a clone that loses its original flexibility. The experimental results generated from such strains reveal the evolutionary adaptations to, and/or non-adaptive changes induced by, the culture conditions under which the strain is maintained.

Intraspecific variations in oxylipin metabolism were also reported for example in *Skeletonema marinoi* (Gerecht *et al.* 2011). This has been discussed in the previous chapter (IV). Thus, a single strain represents only a part of physiological diversity in a species, and when maintained in laboratory for long time may or may not reflect the actual adaptations in a species.

5.4.7 ***Temporal isolation of the species explained by the temperature.*** One of the reasons for undertaking the study was to understand the reason for temporal isolation of the two species in the GoN. *Leptocylindrus danicus* that performs equally well at all Ts, with slightly low μ_{max} at 12°C can be suited to grow throughout the year, Although, the species *L. aporus* that undergoes physiological stress under 12°C but grows equally well at

temperature of 19 and 26°C, is suited to grow in all seasons except winter. However in the natural environment *L. danicus* can be found in all seasons except summer and *L. aporus* can be seen only during summer and autumn (Chapter II). The observation made in the study, that maintenance of long-term cultures of *L. danicus* at 26°C and *L. aporus* at 12°C requires careful monitoring, are in better concordance with what is observed in the natural environment. Thus, from the study, it appears that even a single niche factor like temperature can help answer the temporal isolation of species.

5.4.8 ***Genetic and physiological distances between species.*** The distances of separation of species based on morphology, physiology and biochemical constituents can vary at different levels. The physiological differences observed for *L. danicus* and *L. aporus* in this study point out that the level of diversity seen in morphology can hardly be compared to what is observed in the physiology of the species. The two species being morphologically subtly different responded distinctly to the T conditions exposed. Thus, in the study, there is no significant correlation in the distances inferred from morphology and the distances inferred through the physiological responses. The possible reason for this non-significant correlation might be biased view with a single factor. The differences in physiological traits of the two species are viewed singly to T, but they might sharply vary when combined with other factors especially in the natural environment. Considering this view, morphology and physiology give a different level of estimate of the genetic distance between species.

5.4.9 ***Trade-off in biomass build-up, time and metabolite accumulation.*** The two species show differences in their physiology and metabolite production patterns that point to towards a trade-off between biomass build-up, time and metabolite accumulation. *Leptocylindrus danicus* performs equally well at all three Ts, with slightly low μ_{max} at 12°C, whilst *L. aporus* performs better at the two higher temperatures. Besides, *L. danicus*

accumulates similar biomass in approximately the same time. Instead, *L. aporus* accumulates a higher biomass at a lower temperature, but takes a longer time to reach this high biomass. Therefore, at least in *L. aporus,* there is trade-off between the biomass accumulated and time taken to accumulate the given biomass, which is the ultimate target in any commercial application or sustainability in ecosystem. Furthermore, oxylipin accumulation in *L. danicus* decreases with increase in temperature. If the same trend of *L. danicus* also holds for *L. aporus*, which can accumulate a far higher biomass, then the quantity of metabolites produced constitutes a trade-off between the produced biomass and the time in which this biomass is produced. Thus commercial firms producing secondary metabolites must consider this trade-off in the process of optimising the yield.

5.5 Conclusions

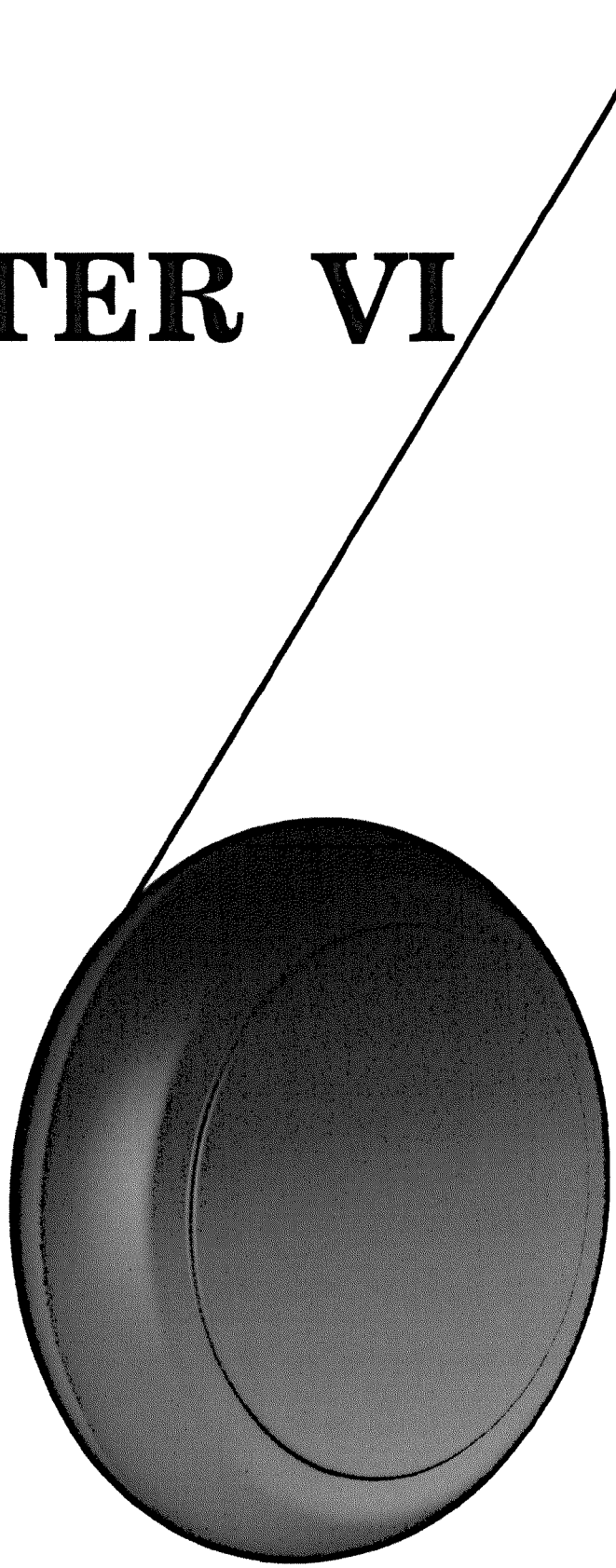
The study helps to improve our understanding of the species adaption to particular environmental condition in the natural environments by providing data on the changes in their cell morphology, growth rates, biomass accumulation and specific biochemical changes to different T. From the study it is evident that abiotic factors like T can profoundly influence the cellular functions causing many physiological and biochemical changes. Responses to the changes in T can be species specific, with each species responding at different levels and exhibiting different patterns. The observed physiological changes especially based on a single factor may or may not reflect the actual distance between the species. Thus, diversity estimates based on physiological factors must consider multiple factors before correlating it with morphological and/or molecular diversity.

Biomass accumulation and pattern of secondary metabolite production are strongly influenced by abiotic factors. Temperature influence on biomass accumulation can be a species specific character.

Temporal distribution of species is largely determined by the environmental factors. Species tend to occur in conditions for which they are evolved and best adapted to grow. In the study, the observation made for *L. danicus* and *L. aporus* in response to different T may explain the contrasting seasonality of the species. Therefore, single abiotic factor like T can provide clues for the temporal isolation of species. Finally studies on individual environmental factors can elicit a response different from that of synergistic effects resulting from multiple factors. Data interpretations on single factor must be done rather carefully as it is rare to have an isolated stress factor in the natural environment.

Thus, a wide variety of information is generated by comprehensive analyses of the physiology and oxylipin quantification at different temperature. The analytic platforms to do metabolite profiling are becoming increasingly powerful by the introduction of Fourier transform-MS and by platforms that combine different LC-MS/MS approaches. Along with the use of increasingly powerful technologies, it is also vital to employ suitable protocols for harvesting, quenching, sample handling, and analysis and to perform laboratory experiments in order to validate that the measurements dependably reflect the levels of metabolites that were present in the cultures.

CHAPTER VI



Conclusions and Perspectives

Marine diatom blooms exhibit recurrent and at times highly predictable patterns (Smayda 1980, Zingone *et al.* 1995). These patterns might be explained by a combination of abiotic and biotic factors. The blooming of a species in different regions and over a broad range of environmental conditions could be explained by differences in the genetic composition among populations in these different regions, or even the existence of different cryptic species in the different regions (Hypothesis 1, see introduction). The alternative hypothesis states that a species blooms in different regions and over a broad range of environmental conditions because each single individual in that population exhibits wide-ranging physiological capabilities, which enables it to cope with different conditions in different environments, i.e., cells possess a high physiological plasticity (Hypothesis 2 in Introduction).

The present studies have shown that these two hypotheses are not mutually exclusive. A single morphological species (sensu lato), believed to bloom all year round in the Gulf of Naples (GoN), was shown to consist of multiple species, each with their own, more confined, blooming period, which may partly overlap with that of other species. The results also demonstrated that these different species exhibit physiological differences as well. All this evidence supports Hypothesis 1. Nonetheless, all the strains tested were able to grow under a wide range of temperature conditions, which provides support for Hypothesis 2.

A test of the first hypothesis of genetic diversity in *Leptocylindrus danicus* (Chapter I) revealed that *Leptocylindrus danicus* consists of multiple genetic units: *L. aporus, L. convexus, L. danicus* and *L. hargravesii*, i.e., four genetically, and ultrastructurally distinct species. Of these, *L. aporus* was considered a variety of *L. danicus,* and *L. convexus* and *L. hargravesii* are new to science. Each of the four species showed a restricted period of occurrence, suggesting that each one of them has a different

seasonality. Thus, the finding of genetically distinct *Leptocylindrus* species within the morphospecies *L. danicus* sensu lato demonstrates that the diversity of a so called cosmopolitan species should be studied in detail and its geographic genetic diversity assessed before drawing inferences about its distribution. The problem with morphological taxonomy of phytoplankton species has been that the majority of phytoplankton species were described for European waters. The resulting taxonomic keys were then forced upon the phytoplankton diversity in other parts of the world.

This study has not only resulted in demonstrating that *L. danicus* sensu lato in the GoN consists of multiple species, it also has raised many new questions. The first question relates to whether geographic sampling increases the number of species detected in a genus such as *Leptocylindrus.*

Leptocylindrus danicus and *L. hargravesii* are closely related because their fast evolving DNA markers (ITS, psbC, rbcL) are highly similar. Peculiarly, the two species occur in sympatry during part of the seasonal cycle in the GoN. This result suggests that these two species do not interbreed. Phenological studies are needed over the seasonal occurrence of the two species to assess if they reproduce sexually in different periods. Mating experiment, using fluorescent tags for differentiating gametes and zygotes in these homothallic species, might resolve if these two species can interbreed successfully. Population genetic studies, using microsatellites, may demonstrate in sharper focus if the two species are indeed entirely separated taxa, or if they interbreed in nature (i.e., if gene flow occurs between the two species). The differences in seasonality observed for the *Leptocylindrus* species raises the question why some species, such as *L. convexus, L. hargravesii,* and *T. belgicus,* exhibit more restricted time periods of occurrence than *L. danicus.* Is this determined by the environmental conditions that the species is adapted to, or is it the internal genetic control mechanisms that control the occurrence of the species.

Although the diversity hypothesis (Hypothesis 1) is basically correct for *L. danicus* sensu lato, it does not necessarily exclude Hypothesis II as the description of species diversity is restricted to a single geographical location. Therefore, prior to testing the other hypothesis it is important to establish the validity and distributions of these species in other geographical locations. Distribution of species can be according to two hypotheses, cosmopolitanism and endemism. As a quick estimate of the spatial distribution of the species (investigated in Chapter I) and to detect additional diversity in diverse geographical locations, metagenomic databases developed at two different geographic scales were utilised: BioMarKs (European) and *Tara* Oceans (worldwide). The diversity inferred from such metagenomic studies depends on the choice of the molecular marker and on the cut-off values used to delineate genetic similarity as a proxy of species adherence. In Chapter III, the V4 and V9 regions of the nuclear SSU rDNA were used to estimate species richness in the natural environment. The BioMarKs V4-based metagenomic database on a European scale did not reveal additional species diversity. Yet it provided additional information about the distribution of the already known and newly described species. Although, the results support the seasonality established at the site LTER-MC, the environmental conditions in which the species can occur varied among sites. For example, the temperature range at which *L. aporus* bloomed in the GoN was distinct from that at which it bloomed in the Oslo Fjord or near Barcelona. *Leptocylindrus aporus* bloomed in Naples and Oslo in the summer season, but under very different temperatures, indicating that temperature cannot be the only factors determining species distribution and seasonality. The *Tara* Oceans dataset based on the V9 marker was generated from samples gathered over a wide geographical area. This dataset provided a larger overview of the distribution of all the *Leptocylindrus* species, and revealed additional diversity within the genus. The species of *Leptocylindrus* identified in Chapter II are not limited to the GoN or Europe, but occur all across the world's oceans, supporting the hypothesis of cosmopolitanism for

these species. Notably, *L. minimus* was not detected in the *Tara* database, possibly because the Tara Oceans expedition did not sequence samples from Northern cold waters (as of the time of sequence retrieval for this study). Additional clades, possibly belonging to yet unknown *Leptocylindrus* species were found to be restricted to smaller geographical regions or even single sample sites, which in contrast, challenges the hypothesis of cosmopolitism. This finding is similar to what was observed for some of the *Skeletonema* species in Kooistra et al. (2008). In any case, the results of the present study suggest that species diversity recovered in the environment increases with the sampling scale. This first attempt to reconstruct species distribution using metagenomic databases might have been influenced by the choice of molecular maker and the cut-off values used to consider the genetic identity. Additionally, the methods developed in the study can be replicated with other known diatoms species or even other diverse species. The study is one of the first attempts studying the spatial distribution of individual species of microalgae worldwide, while similar attempts have been made in microbial system.

Through the application of metagenomic databases the species discovered in Chapter II were established as to be present not just in the GoN but also in Europe and other parts of the world. *Leptocylindrus minimus* that was absent in the samples of GoN was recovered from the Oslo Fjord sequences. There are many more questions to be answered. Why the species *L. minimus* is present only in the North Sea and not also in the cold waters of the Southern Ocean such as the Antarctic Peninsula (sampled by Tara Oceans)? It might be that the species was absent during the time of sampling. Multiple sampling at a single location across different seasons is proposed to show that species can be recovered at a single site. However, the repeated sampling at the GoN has not revealed the presence of the species. So, the question why this species is restricted to the North Sea remains to be answered. The phenology of *Tenuicylindrus belgicus* differs from that of *L. minimus*, in that the former species blooms exclusively in late summer in

GoN (that is, in warm water), but instead in spring in Oslo (that is, in cold water). This situation is similar to that observed in *Chaetoceros socialis* (Degerlund *et al.* 2012). Is the latter species the same at both sites? Why does it have a narrow blooming time? What are the physiological adaptations in the strains of these distinct geographical locations that cause the differences in the seasonality? Likewise each species has its own unique characteristics that pose many questions. In addition the general questions, again, remain to be answered. Does species diversity increase with multiple location sampling? Is everything everywhere, or are some species everywhere and others not? If the latter is correct, then why are some species cosmopolitan whereas others have restricted distribution patterns? Do the cosmopolitans differ in their physiological mechanisms from the regionalists? I tried to develop and pursue this question in the next two chapters

Although morphological and DNA-based markers are regularly applied for species identification and delineation, alternative methods including biochemical markers can be applied especially for species that are closely related. LOX-derived polyunsaturated fatty acid profiles differed markedly amongst the species tested, whereas they showed only very minor differences among strains belonging to the same species. Oxylipin profiles provided the same resolution as DNA markers and morphology. In fact, even the closely related species, *L. danicus* and *L. hargravesii* could be differentiated using their oxylipin profiles. Thus, biochemical markers could be used to discriminate among cryptic species, if not exclusively in combination with the morphological and DNA based markers.

Oxylipin pathways are widespread in photosynthetic organisms. Their application as biomarkers still remains to be explored. Some of the strains, such as *L. aporus* SZN-B651, do not produce oxylipins. To understand why these strains do not produce oxylipins, their transcriptome might provide indications why this is so. Moreover,

whole transcriptomes of different strains can be compared to explain metabolomic differences between strains.

The second hypothesis of the study, physiological diversity, was also included and tested. Temperature, one of the major determining factors of species occurrence was used to understand the physiological diversity in the species *L. danicus* and *L. aporus*. The two species occur in different seasons in the GoN and, thus, can reflect differential physiological adaptation. *Leptocylindrus danicus* is present in all seasons except in summer. Results (Chapter V) show that it is unable to withstand long-term exposure to 26 °C, which is a temperature typical for the summer conditions of the surface water in the GoN. *Leptocylindrus aporus*, the species responsible for intense bloom in summer in the GoN, can grow under these high temperature conditions. Instead, *L. aporus* could not withstand long-term exposure to 12 °C, whereas *L. danicus* can. Overall, these results indicate that *L. danicus* can better sustain low temperatures whereas *L. aporus* can cope better with high temperatures. However, the seasonal restriction observed for the species are based on the strains grown under laboratory conditions. Environmental sequencing results of the two species at the GoN demonstrates the presence of these species in the environment during the season in which they do not bloom, but apparently exist in low numbers that are difficult to detect by traditional methods. Nevertheless, the observed pattern of seasonality in the natural environment corroborates with the physiological responses in situ. Thus, physiological adaption determines the species occurrence in diverse environments. An alternative view is that the species phased on a selected season by internal, genetic time control eventually lose their capability to withstand different conditions. This mechanism, i.e. removing adaptations to conditions not really experienced, probably involve some energetic advantages or at least does not have an impact on the fitness of the species in that environment, although it reduces its plasticity. In this case, the selectivity for certain

temperature ranges could be the consequences and not the cause for the observed seasonality.

The analysis presented in Chapter V is the first step for future studies that are planned on this material. Addressing the whole metabolite pattern (metabolome) will reveal metabolic changes that occur at three temperatures in *L. aporus* and *L. danicus*, thereby providing an explanation for the observed physiological differences. This explanation has important implications for the biotechnology industries where maximum product yield is the ultimate aim. An optimised growth condition is mandatory for balancing the maximal biomass yield and obtaining maximal product of interest.

Overall, the presence of multiple species within a morphologically simple species, their seasonal distribution, the differences in special distribution, and differential responses to abiotic factors, all suggests that genetically distinct species may play different ecological and biogeochemical roles in marine habitats. The succession of genetically distinct species over the year under distinct environmental conditions (Chapter II) suggests a close connection between the genetics and the environment in which a species persists. The difference in the spatial distribution, with restrictions of some genotypes to a geographical location and others to multiple locations (Chapter III), again strengthens the connection between the genetics and the environment in which a species persists, suggesting that the two hypotheses of species biogeography are valid, i.e., not mutually exclusive. The physiological adaptation in closely related species, with each species responding differently to variation in abiotic factors, adds further strength to the connection between the genetics and the environment, suggesting an environmental role in shaping community structure. Understanding this connection might help better explain, and perhaps predict, blooms of individual species, which is particularly interesting for invasive and harmful algal-bloom species, but also for predicting alterations of species distribution which in turn may alter biogeochemical cycles under the ongoing and predicted climate change

Bibliography

AL-KUBAISI, K.H., SCHWANTES, H. O. 1981. Cytophotometrische Untersuchungen zum Generationswechsel autotropher and heterotropher siphonaler Organismen (*Vaucheria sessilis* and *Saprolegnia ferax*). *Nova Hedwigia* 34: 301-316.

ALLEN, A.E., VARDI, A. & BOWLER, C. 2006. An ecological and evolutionary context for integrated nitrogen metabolism and related signaling pathways in marine diatoms. *Current Opinion in Plant Biology* 9: 264-273.

ALVERSON, A.J. 2008. Molecular systematics and the diatom species. *Protist* 159: 339-353.

AMARAL-ZETTLER, L.A., MCCLIMENT, E.A., DUCKLOW, H.W. & HUSE, S.M. 2009. A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA Genes. *PLoS One* 4: e6372.

AMATO, A., KOOISTRA, W.H.C.F., LEVIALDI GHIRON, J.H., MANN, D.G., PRÖSCHOLD, T. & MONTRESOR, M. 2007. Reproductive isolation among sympatric cryptic species in marine diatoms. *Protist* 158: 193-207.

AMATO, A. & MONTRESOR, M. 2008. Morphology, phylogeny, and sexual cycle of *Pseudo-nitzschia mannii* sp. nov. (Bacillariophyceae): a pseudo-cryptic species within the *P. pseudodelicatissima* complex. *Phycologia* 47: 487–497.

ANDREOU, A., BRODHUN, F. & FEUSSNER, I. 2009. Biosynthesis of oxylipins in non-mammals. *Progress in Lipid Research* 48: 148-170.

BAAS-BECKING, L.G.M. 1934. *Geobiologie of inleiding tot de milieukunde.* Den Haag : W.P. Van Stockum & Zoon. , The Hague. 263 pp.

BALZANO, S., SARNO, D. & KOOISTRA, W.H.C.F. 2011. Effects of salinity on the growth rate and morphology of ten *Skeletonema* strains. *Journal of Plankton Research* 33: 937-945

BAO, E., JIANG, T., KALOSHIAN, I. & GIRKE, T. 2011. SEED: efficient clustering of next-generation sequences. *Bioinformatics* 27: 2502-2509.

BARACCHI, D., DAPPORTO, L., TESEO, S., HASHIM, R. & TURILLAZZI, S. 2010. Medium molecular weight polar substances of the cuticle as tools in the study of the taxonomy, systematics and chemical ecology of tropical hover wasps (Hymenoptera: Stenogastrinae). *Journal of Zoological Systematics and Evolutionary Research* 48: 109-114.

BERRY, J.A. & RAISON, J.K. 1981. Responses of macrophytes to temperature. *Encyclopedia of plant physiology, New series* 12: 277-338.

BESZTERI, B., JOHN, U. & MEDLIN, L.K. 2007. An assessment of cryptic genetic diversity within the *Cyclotella meneghiniana* species complex (Bacillariophyta) based on nuclear and plastid genes, and amplified fragment length polymorphisms. *European Journal of Phycology* 42: 47-60.

BOERO, F. 1996. Episodic events: their relevance to ecology and evolution. *P.S.Z.N.I: Marine Ecology* 17: 237-250.

BÖLLING, C. & FIEHN, O. 2005. Metabolite profiling of *Chlamydomonas reinhardtii* under nutrient deprivation. *Plant Physiology* 139: 1995-2005.

BOSE, A., TIWARI, B., CHATTOPADHYAY, M.K., GUPTA, S. & GHOSH, B. 1999. Thermal stress induces differential degradation of RuBisCo in heat-sensitive and heat-tolerant rice. *Physiologia Plantarum* 105: 89-94.

BOYD, P.W., LAW, C.S., WONG, C., NOJIRI, Y., TSUDA, A., LEVASSEUR, M., TAKEDA, S., RIVKIN, R., HARRISON, P.J., STRZEPEK, R., GOWER, J., MCKAY, R.M., ABRAHAM, E., ARYCHUK, M., BARWELL-CLARKE, J., CRAWFORD, W., CRAWFORD, D., HALE, M., HARADA, K., JOHNSON, K., KIYOSAWA, H., KUDO, I., MARCHETTI, A., MILLER, W., NEEDOBA, J., NISHIOKA, J., OGAWA, H., PAGE, J., ROBERT, M., SAITO, H., SASTRI, A., SHERRY, N., SOUTAR, T., SUTHERLAND, N., TAIRA, Y., WHITNEY, F., WONG, S.-K.E. & YOSHIMURA, T. 2004. The decline and fate of an iron-induced subarctic phytoplankton bloom *Nature* 428: 549-553.

BRAARUD, T. 1962. Species distribution in marine phytoplankton. *Journal of the Oceanographical Society of Japan* 20th Anniversary Volume: 628-649.

CALDWELL, G.S., OLIVE, P.J.W. & BENTLEY, M.G. 2002. Inhibition of embryonic development and fertilization in broadcast spawning marine invertebrates by water soluble diatom extracts and the diatom toxin 2-trans, 4-trans decadienal. *Aquatic Toxicology* 60: 123-137.

CALDWELL, G.S., BENTLEY, M.G. & OLIVE, P.J.W. 2004. First evidence of sperm motility inhibition by the diatom aldehyde 2E, 4E-decadienal. *Marine Ecology Progress Series* 273: 97-108.

CARON, D.A., COUNTWAY, P.D., SAVAI, P., GAST, R.J., SCHNETZER, A., MOORTHI, S.D., DENNETT, M.R., MORAN, D.M. & JONES, A.C. 2009. Defining DNA-based operational taxonomic units for microbial-eukaryote ecology. *Applied and Environmental Microbiology* 75: 5797-5808.

CASE, R.J., BOUCHER, Y., DAHLLÖF, I., HOLMSTRÖM, C., DOOLITTLE, W.F. & KJELLEBERG, S. 2007. Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. *Applied and Environmental Microbiology* 73: 278-288.

CASOTTI, R., MAZZA, S., BRUNET, C., VANTREPOTTE, V., IANORA, A. & MIRALTO, A. 2005. Growth inhibition and toxicity of the diatom aldehyde 2-trans, 4-trans-decadienal on *Thalssiosira weissflogii* (Bacillariophyceae). *Journal of Phycology* 41: 7-20.

CAVALIER-SMITH, T. 1993. The origin, losses and gains of chloroplasts. In *Origin of Plastids.* (Ed. by R.A. Lewin), pp. 291-348. Chapman & Hall, New York.

CAVALIER-SMITH, T. 2000. Membrane heredity and early chloroplast evolution. *Trends in Plant Science* 5: 174-182.

CAVENDER-BARES, K., CHISHOLM, S., ONDRUSEK, M. & BIDIGARE, R. 1999. Differential response of equatorial Pacific phytoplankton to iron fertilization. *Limnology and Oceanography* 44: 237-246.

CAVENDER-BARES, K.K., KARL, D.M. & CHISHOLM, S.W. 2001. Nutrient gradients in the western North Atlantic Ocean: Relationship to microbial community structure and comparison to patterns in the Pacific Ocean. *Deep Sea Research Part I: Oceanographic Research Papers* 48: 2373-2395.

CERMEÑO, P. & FALKOWSKI, P.G. 2009. Controls on diatom biogeography in the ocean. *Science* 325: 1539-1541.

CHEPURNOV, V.A. & MANN, D.G. 1997. Variation in the sexual behaviour of natural clones of *Achnanthes longipes* (Bacillariophyta). *European Journal of Phycology* 32: 147-154.

CHEPURNOV, V.A. & MANN, D.G. 1999. Variation in the sexual behaviour of *Achnanthes longipes* (Bacillariophyta). II. Inbred monoecious lineages. *European Journal of Phycology* 34: 1-11.

CHEPURNOV, V.A., MANN, D.G., SABBE, K. & VYVERMAN, W. 2004. Experimental studies on sexual reproduction in diatoms. *International Review of Cytology* 237: 91-154.

CLEMENT, A. & LEMBEYE, G. 1993. Phytoplankton monitoring program in the fish farm region of the South of Chile. In *Toxic phytoplankton blooms in the sea.* (Ed. by T.J. Smayda & Y. Shimizu), pp. 223-228. Elsevier, Amsterdam.

CLEVE, P.T. 1889. Pelagisk diatomeer fran Kattegat. In *Det èvidenskabelige udbytte af kanonbaaden "Hauchs" togter i de Danske have indefor Skagen, I. Aarene.* (Ed. by C.G.J. Petersen), pp. 53-56. Andr. Fred. Høst & Sons Forlag., Kjøbenhavn.

CLEVE, P.T. 1894. Redogorelse for de Sevnska hydrografiska undersokningarne aren 1893-1894. In *II. Planktonundersökninoar.* (Ed. by G. Ekman, O. Pettersson & A. Wijkander), pp. 1-16. P. A. Nordstedt & Söner Stockholm.

CLOERN, J.E. 1996. Phytoplankton bloom dynamics in coastal ecosystems: a review with some general lessons from sustained investigation of San Francisco Bay, California. *Reviews of Geophysics* 34: 127-168.

COLEMAN, A.W. 2002. Microbial eukaryote species. *Science* 297: 337.

COOK, C.M., VARDAKA, E. & LANARAS, T. 2004. Toxic cyanobacteria in Greek freshwaters, 1987-2000: Occurrence, toxicity, and impacts in the Mediterranean region. *Acta Hydrochimica et Hydrobiologica* 32: 107-124.

CRAWFORD, R.M. & SIMS, P.A. 2008. Some principles of chain formation as evidenced by the early fossil record. *Nova Hedwigia Beiheft* 133: 171-176.

CUTIGNANO, A., LAMARI, N., D'IPPOLITO, G., MANZO, E., CIMINO, G. & FONTANA, A. 2011. Lipoxygenase products in marine diatoms: a concise analytical method to explore the functional potential of oxylipins. *Journal of Phycology* 47: 233-243.

D'IPPOLITO, G., IADICICCO, O., ROMANO, G. & FONTANA, A. 2002. Detection of short-chain aldehydes in marine organisms: the diatom *Thalassiosira rotula*. *Tetrahedron Letters* 43: 6137-6140.

D'IPPOLITO, G., TUCCI, S., CUTIGNANO, A., ROMANO, G., CIMINO, G., MIRALTO, A. & FONTANA, A. 2004. The role of complex lipids in the synthesis of bioactive aldehydes of the marine diatom *Skeletonema costatum*. *Biochimica et Biophysica Acta* 1686: 100-107.

D'IPPOLITO, G., LAMARI, N., MONTRESOR, M., ROMANO, G., CUTIGNANO, A., GERECHT, A., CIMINO, G. & FONTANA, A. 2009. 15S-Lipoxygenase metabolism in the marine diatom *Pseudo-nitzschia delicatissima*. *New Phytologist* 183: 1064-1071

D'IPPOLITO, G., LAMARI, N., MONTRESOR, M., ROMANO, G., CUTIGNANO, A., GERECHT, A., CIMINO, G. & FONTANA, A. 2009. 15S-Lipoxygenase metabolism in the marine diatom Pseudo-nitzschia delicatissima. *New Phytologist* 183: 1064-1071.

DAVIS, C.O., HOLLIBAUGH, J.T., SEIBERT, D.L.R., THOMAS, W.H. & HARRISON, P.J. 1980. Formation of resting spores by *Leptocylindrus danicus* (Bacillariophyceae) in a controlled experimental ecosystem. *Journal of Phycology* 16: 296-302.

DE CASTRO ARAÚJO, S. & GARCIA, V.M.T. 2005. Growth and biochemical composition of the diatom *Chaetoceros* cf. *wighamii* Brightwell under different temperature, salinity and carbon dioxide levels. I. Protein, carbohydrates and lipids. *Aquaculture* 246: 405-412.

DE MARTINO, A., MEICHENIN, A., SHI, J., PAN, K. & BOWLER, C. 2007. Genetic and phenotypic characterization of *Phaeodactylum tricornutum* (Bacillariophyceae) accessions. *Journal of Phycology* 43: 992 - 1009.

DEGERLUND, M., HUSEBY, S., ZINGONE, A., SARNO, D. & LANDFALD, B. 2012. Functional diversity in cryptic species of *Chaetoceros socialis* Lauder (Bacillariophyceae). *Journal of Plankton Research* 34: 416-431.

DETTMER, K., ARONOV, P.A. & HAMMOCK, B.D. 2007. Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews* 26: 51-78.

DEVOS, N., INGOUFF, M., LOPPES, R. & MATAGNE, R.F. 1998. RuBisCo adaptation to low temperatures: a comparative study in psychrophilic and mesophilic unicellular algae. *Journal of Phycology* 34: 655-660.

DICKSON, D.M.J. & KIRST, G.O. 1987. Osmotic adjustment in marine eukariotic algae: the role of inorganic ions, quaternary ammonium, tertiary sulphonium and carbohydrate solutes. *New Phytologist* 106: 645-655.

DIERSING, N. 2009. Phytoplankton blooms: the basics. *National Oceanic and Atmospheric Administration*: 1-2.

DINSDALE, E., EDWARDS, R., HALL, D., ANGLY, F., BREITBART, M., BRULC, J., FURLAN, M., DESNUES, C., HAYNES, M., LI, L., MCDANIEL, L., MORAN, M., NELSON, K., NILSSON, C., OLSON, R., PAUL, J., BRITO, B., RUAN, Y., SWAN, B., STEVENS, R., VALENTINE, D., THURBER, R., WEGLEY, L., WHITE, B. & ROHWER, F. 2008. Functional metagenomic profiling of nine biomes. *Nature* 452: 629-632.

DUNSTAN, G.A., VOLKMAN, J.K., BARRETT, S.M., LEROI, J.-M. & JEFFREY, S.W. 1993. Essential polyunsaturated fatty acids from 14 species of diatom (Bacillariophyceae). *Phytochemistry* 35: 155-161.

ECKARDT, N.A. 2008. Oxylipin signaling in plant stress responses. *The Plant Cell Online* 20: 495-497.

EDGAR, R.C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26: 2460-2461.

EILERTSEN, H.C.H.R., SANDBERG, S. & TOLLEFSEN, H. 1995. Photoperiodic control of diatom spore growth: a theory to explain the onset of phytoplankton blooms. *Marine Ecology Progress Series* 116: 303-307.

EPPLEY, R.W. 1972. Temperature and phytoplankton growth in the sea. *Fishery Bulletin* 70: 1063–1085.

EVANS, K.M., WORTLEY, A.H., SIMPSON, G.E., CHEPURNOV, V.A. & MANN, D.G. 2008. A molecular systematic approach to explore diversity within the *Sellaphora pupula* species complex (Balcillariophyta). *Journal of Phycology* 44: 215-231.

FALKOWSKI, P.G. & RAVEN, J.A. (eds.) 1997. *Aquatic photosynthesis.* Blackwell science, Massachusetts, USA.

FENCHEL, T., ESTEBAN, G.F. & FINLAY, B.J. 1997. Local *versus* global diversity of microorganisms: cryptic diversity of ciliated protozoa. *Oikos* 80: 220-225.

FENCHEL, T. & FINLAY, B.J. 2006. The diversity of microbes: resurgence of the phenotype. *Philosophical Transactions of the Royal Society B: Biological Sciences* 361: 1965-1973.

FENCHEL, T.O.M. & FINLAY, B.J. 2004. The ubiquity of small species: patterns of local and global diversity. *Bioscience* 54: 777-784.

FIALA, M. & ORIOL, L. 1990. Light-temperature interactions on the growth of Antarctic diatoms. *Polar Biology* 10: 629-636.

FIDALGO, J.P., CID, A., TORRES, E., SUKENIK, A. & HERRERO, C. 1998. Effects of nitrogen source and growth phase on proximate biochemical composition, lipid classes and fatty acid profile of the marine microalga *Isochrysis galbana. Aquaculture* 166: 105-116.

FIELD, C., BEHRENFELD, M., RANDERSON, J. & FALKOWSKI, P. 1998. Primary production of the biosphere: integrating terrestrial and oceanic components *Science* 281: 237 - 240.

FINKEL, Z.V. & IRWIN, A.J. 2000. Modeling size-dependent photosynthesis: Light absorption and the allometric rule. *Journal of Theoretical Biology* 204: 361-369.

FINLAY, B.J. & CLARKE, K.J. 1999. Ubiquitous dispersal of microbial species. *Nature* 400: 828.

FINLAY, B.J. 2002. Global dispersal of free-living microbial eukaryote species. *Science* 296: 1061-1063.

FONTANA, A., D'IPPOLITO, G., CUTIGNANO, A., ROMANO, G., LAMARI, N., GALLUCCI, A.M., CIMINO, G., MIRALTO, A. & IANORA, A. 2007. LOX-induced lipid peroxidation mechanism responsible for the detrimental effect of marine diatoms on zooplankton grazers. *Chembiochem* 8: 1810-1818.

FRENCH, F. & HARGRAVES, P.E. 1980. Physiological characteristics of plankton diatom resting spores. *Marine Biology Letters* 1: 185-195.

FRENCH III, F.W. & HARGRAVES, P.E. 1985. Spore formation in the life cycles of the diatoms *Chaetoceros diadema* and *Leptocylindrus danicus*. *Journal of Phycology* 21: 477-483.

FRENCH III, F.W. & HARGRAVES, P.E. 1986. Population dynamics of the spore-forming diatom *Leptocylindrus danicus* in Narragansett Bay, Rhode Island. *Journal of Phycology* 22: 411-420.

FURNAS, M.J. 1990. *In situ* growth rates of marine phytoplankton: approaches to measurement, community and species growth rates. *Journal of Plankton Research* 21: 1117-1151.

GALLAGHER, J.C. 1983. Cell enlargement in *Skeletonema costatum* (Bacillariophyceae). *Journal of Phycology* 19: 539-542.

GEIDER, R.J., PLATT, T. & RAVEN, J.A. 1986. Size dependence of growth and photosynthesis in diatom: a synthesis. *Marine Ecology Progress Series* 30: 93-104.

GENOVESI-GIUNTI, B., LAABIR, M. & VAQUER, A. 2006. *The benthic resting cyst : A key actor in harmful dinoflagellate blooms -A review.* Laboratoire Arago, Université Pierre et Marie Curie, Banyuls-sur-Mer. 11 pp.

GERECHT, A., ROMANO, G., IANORA, A., D'IPPOLITO, G., CUTIGNANO, A. & FONTANA, A. 2011. Plasticity of oxylipin metabolism among of the marine diatom *Skeletonema marinoi* (Bacillariophyceae) *Journal of Phycology* 47: 1050-1056.

GERWICK, W.H., MOGHADDAM, M. & HAMBERG, M. 1991. Oxylipin metabolism in the red alga *Gracilariopsis lemaneiformis*: Mechanism of formation of vicinal dihydroxy fatty acids. *Archives of Biochemistry and Biophysics* 290: 436-444.

GHODSI, M., LIU, B. & POP, M. 2011. DNACLUST: accurate and efficient clustering of phylogenetic marker genes. *BMC Bioinformatics* 12: 271.

GIBBS, S.P. 1979. The route of entry of cytoplasmically synthesized proteins into chloroplasts of algae possessing chloroplast ER. *Journal of Cell Science* 35: 253-266.

GIBBS, S.P. 1981. The chloroplasts of some algal groups may have evolved from endosymbiotic eukaryotic algae. *Annals of the New York Academy of Sciences* 361: 193-208.

GMELIN, J.F. 1791. *Carolia Linneaus, Systema naturae per regna tria naturae, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis.* Lipsiae : impensis Georg. Emanuel. Beer, Leipzig 3021-3910 pp.

GOMEZ, A. 2007. The consortium of the protozoan *Solenicola setigera* and the diatom *Leptocylindrus mediterraneus* in the Pacific Ocean. *Acta Protozoologica* 46: 15-24.

GOMEZ, F. 2009. *Torodinium* and *Pavillardia* (Gymnodiniales, Dinophyceae): two unarmoured dinoflagellates with a body extension, collected from the open Pacific Ocean. *Protistology* 6: 131-135.

GORDON, R., LOSIC, D., TIFFANY, M.A., NAGY, S.S. & STERRENBURG, F.A.S. 2009. The glass menagerie: diatoms for novel applications in nanotechnology. *Trends in Biotechnology* 27: 116-127.

GRAN, H.H. 1915. *The plankton production in the north European waters in the spring of 1912.* Bureau du Conseil permanent international pour l'exploration de la mer, Paris. 142 pp.

GREEN, J. & BOHANNAN, B.J.M. 2006. Spatial scaling of microbial biodiversity. *Trends in Ecology & Evolution* 21: 501-507.

GROUP, T.N.H.W., PETERSON, J., GARGES, S., GIOVANNI, M., MCINNES, P., WANG, L., SCHLOSS, J.A., BONAZZI, V., MCEWEN, J.E., WETTERSTRAND, K.A., DEAL, C., BAKER, C.C., DI FRANCESCO, V., HOWCROFT, T.K., KARP, R.W., LUNSFORD, R.D., WELLINGTON, C.R., BELACHEW, T., WRIGHT, M., GIBLIN, C., DAVID, H., MILLS, M., SALOMON, R., MULLINS, C., AKOLKAR, B., BEGG, L., DAVIS, C., GRANDISON, L., HUMBLE, M., KHALSA, J., LITTLE, A.R., PEAVY, H., PONTZER, C., PORTNOY, M., SAYRE, M.H., STARKE-REED, P., ZAKHARI, S., READ, J., WATSON, B. & GUYER, M. 2009. The NIH human microbiome project. *Genome Research* 19: 2317-2323.

GUILLARD, R.L. & RYTHER, J.H. 1962. Studies of marine planktonic diatoms. I. Cyclotella nana Hustedt, and Detonula confervacea (Cleve) Gran. *Canadian Journal of Microbiology* 8: 229-239.

HAECKEL, E. 1873. Ueber einige neue pelagische Infusorien. In *Ibid.,*. pp. 561-7 pp., 27,28 pls.

HALLEGRAEFF, G.M., ANDERSON, D.M., CEMBELLA, A.D., ENEVOLDSEN, H.O. & UNESCO 2003. *Manual on harmful marine microalgae.* Unesco. pp.

HALSBAND-LENK, C., PIERSON, J.J. & LEISING, A.W. 2005. Reproduction of *Pseudocalanus newmani* (Copepoda: Calanoida) is deleteriously affected by diatom blooms – A field study. *Progress In Oceanography* 67: 332-348.

HANNAH, M.A., CALDANA, C., STEINHAUSER, D., BALBO, I., FERNIE, A.R. & WILLMITZER, L. 2010. Combined transcript and metabolite profiling of *Arabidopsis* grown under widely variant growth conditions facilitates the identification of novel metabolite-mediated regulation of gene expression. *Plant Physiology* 152: 2120-2129.

HANSSON, L.A. 1995. Diurnal recruitment patterns in algae: effects of light cycles and stratified conditions. *Journal of Phycology* 31: 540-546.

HARGRAVES, P.E. 1976. Studies on marine plankton diatoms. II. Resting spores morphology. *Journal of Phycology* 12: 118-128.

HARGRAVES, P.E. 1990. Studies on marine plankton diatoms. V. Morphology and distribution on *Leptocylindrus minimus* Gran. *Nova Hedwigia Beiheft* 100: 47-60.

HARRIS, G.P. 1986. *Phytoplankton ecology. Structure, function and fluctuation.* Chapman, London. 1-434 pp.

HARWOOD, J.L. 1988. Fatty acid metabolism. *Annual Review of Plant Physiology and Plant Molecular Biology* 39: 101-138.

HASLE, G.R. 1975. Some living marine species of the diatom family Rhizosoleniaceae. *Nova Hedwigia Beiheft* 39: 99-140.

HASLE, G.R. 1976. The biogeography of some marine planktonic diatoms. *Deep Sea Research* 23: 319-338.

HASLE, G.R. 2001. The marine, planktonic diatom family Thalassionemataceae: morphology, taxonomy and distribution. *Diatom Research* 16: 1-82.

HAY, M.E. 1996. Marine chemical ecology: what's known and what'next?. *Journal of Experimental Marine Biology and Ecology* 200: 103-134.

HAY, M.E. 2002. Preface: the next wave in aquatic chemical ecology. *Journal of Chemical Ecology* 28: 1897-1899.

HAYAKAWA, T., KUDOH, S., SUZUKI, Y. & TAKAHASHI, M. 1994. Temperature-dependent changes in colony size of the freshwater pennate diatom *Asterionella formosa* (Bacillariophyceae) and their possible implications. *Journal of Phycology* 30: 955-964.

HENDEY, N.I. 1964. *An introductory account of the smaller algae of British coastal waters. V. Bacillariophyceae (Diatoms).* Her Majesty's Stationery Office, London. 317 pp.

HILLIS, D.M. & HUELSENBECK, J.P. 1992. Signal, noise, and reliability in molecular phylogenetic analyses. *Journal of Heredity* 83: 189-195.

HILLS, D.M., MORITZ, C. & MABLE, B.K. 1996. *Molecular systematics.* Sinauer Associates, Sunderland. 655 pp.

HU, H. & GAO, K. 2006. Response of growth and fatty acid compositions of *Nannochloropsis* sp. to environmental factors under elevated CO2 concentration. *Biotechnology Letters* 28: 987-992.

HUANG, Y., NIU, B., GAO, Y., FU, L. & LI, W. 2010. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* 26: 680-682.

HUFLEJT, M.E., TREMOLIERES, A., PINEAU, B., LANG, J.K., HATHEWAY, J. & PACKER, L. 1990. Changes in membrane lipid composition during saline growth of the fresh water cyanobacterium *Synechococcus* 6311. *Plant Physiology* 94: 1512-1521.

HUSE, S.M., WELCH, D.M., MORRISON, H.G. & SOGIN, M.L. 2010. Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environmental Microbiology* 12: 1889-1898.

HUSTEDT, F. 1962. *Die kieselalgen. Deutschlands, österrreichs und der schweiz under berücksichtigung der übrigen länder europas sowie der angrenzenden Meeresgebiete. 3. Teil.* Akademische Verlagsgesellschaft m.b.h., Leipzig. 816 pp.

IANORA, A. & POULET, S.A. 1993. Egg viability in the copepod *Temora stylifera*. *Limnology and Oceanography* 38: 1615-1626.

IANORA, A., MIRALTO, A., POULET, S.A., CAROTENUTO, Y., BUTTINO, I., ROMANO, G., CASOTTI, R., POHNERT, G., WICHARD, T., COLUCCI-D'AMATO, L., TERRAZZANO, G. & SMETACEK, V. 2004. Aldehyde suppression of copepod recruitment in blooms of a ubiquitous planktonic diatom. *Nature* 429: 403-407.

JØRGENSEN, E.G. 1968. The Adaptation of Plankton Algae. *Physiologia Plantarum* 21: 423-427.

KARENTZ, D. & SMAYDA, T.J. 1984. Temperature and seasonal occurrence patterns of 30 dominant phytoplankton species in Narragansett Bay over a 22-year period (1959-1980). *Marine Ecology Progress Series* 18: 277-293.

KARENTZ, D. & SMAYDA, T.J. 1998. Temporal patterns and variations in phytoplankton community organization and abundance in Naragansett Bay during 1959-1980. *Journal of Plankton Research* 20: 145-168.

KARSENTI, E., ACINAS, S.G., BORK, P., BOWLER, C., DE VARGAS, C., RAES, J., SULLIVAN, M., ARENDT, D., BENZONI, F., CLAVERIE, J.-M., FOLLOWS, M., GORSKY, G., HINGAMP, P., IUDICONE, D., JAILLON, O., KANDELS-LEWIS, S., KRZIC, U., NOT, F., OGATA, H., PESANT, S., REYNAUD, E.G., SARDET, C., SIERACKI, M.E., SPEICH, S., VELAYOUDON, D., WEISSENBACH, J., WINCKER, P. & TARA OCEANS, C. 2011. A holistic approach to marine eco-systems biology. *PLoS Biology* 9.

KARSTEN, U., WIENCKE, C. & KIRST, G.O. 1992. Dimethylsulphoniopropionate (DMSP) accumulation in green macroalgae from polar to temperate regions: interactive effects of light *versus* salinity and light *versus* temperature. *Polar Biology* 12: 603-607.

KELLER, M.D., SELVIN, R.C., CLAUS, W. & GUILLARD, R.R.L. 1987. Media for the culture of oceanic ultraphytoplankton. *Journal of Phycology* 23: 633-638.

KHOZIN-GOLDBERG, I. & COHEN, Z. 2006. The effect of phosphate starvation on the lipid and fatty acid composition of the fresh water eustigmatophyte *Monodus subterraneus*. *Phytochemistry* 67: 696-701.

KIMURA, M. 1968. Evolutionary rate at the molecular level. *Nature* 217: 624-626.

KING, J.L. & JUKES, T.H. 1969. Non-Darwinian Evolution. *Science* 164: 788-798.

KIRK, J.T.O. 1994. *Light and photosynthesis in aquatic ecosystems.* Cambridge University Press, Cambridge. 528 pp.

KNEITEL, J.M. & CHASE, J.M. 2004. Trade-offs in community ecology: linking spatial scales and species coexistence. *Ecology Letters* 7: 69-80.

KNIGHT, V.I., WANG, H., LINCOLN, J.E., LULAI, E.C., GILCHRIST, D.G. & BOSTOCK, R.M. 2001. Hydroperoxides of fatty acids induce programmed cell death in tomato protoplasts. *Physiological and Molecular Plant Pathology* 59: 277-286.

KOOISTRA, W.H.C.F., DE STEFANO, M., MEDLIN, L.K. & MANN, D.G. 2003. The phylogenetic position of *Toxarium*, a pennate-like lineage within centric diatoms (Bacillariophyceae). *Journal of Phycology* 39: 185-197.

KOOISTRA, W.H.C.F., GERSONDE, R., MEDLIN, L.K. & MANN, D.G. 2007. The origin and evolution of the diatoms: their adaptation to a planktonic existence. In *Evolution of Primary Producers in the Sea* (Ed. by P. Falkowski & A.H. Knoll), pp. 207-250. Elsevier Academic Press, Burlington.

KOOISTRA, W.H.C.F., SARNO, D., BALZANO, S., GU, H., ANDERSEN, R.A. & ZINGONE, A. 2008. Global diversity and biogeography of *Skeletonema* species (Bacillariophyta). *Protist* 159: 177-193.

KOOISTRA, W.H.C.F., SARNO, D., HERNÁNDEZ-BECERRIL, D.U., ASSMY, P., DI PRISCO, C. & MONTRESOR, M. 2010. Comparative molecular and morphological phylogenetic analyses of taxa in the Chaetocerotaceae (Bacillariophyta). *Phycologia* 49: 471-500.

KRABERG, A., BAUMANN, M. & DÜRSELEN, C.D. 2010. *Coastal phytoplankton: photo guide for Northern European seas.* Verlag Dr. Friedich Pfeil, München. 204 pp.

KRÄBS, G. & BÜCHEL, C. 2011. Temperature and salinity tolerances of geographically separated *Phaeodactylum tricornutum* Böhlin strains: maximum quantum yield of primary photochemistry, pigmentation, proline content and growth. *Botanica Marina* 54: 231.

KUDO, I., MIYAMOTO, M., NOIRI, Y. & MAITA, Y. 2000. Combined effects of temperature and iron on the growth and physiology of the marine diatom *Phaedactilum tricornutum* (Bacillariophyceae). *Journal of Phycology* 36: 1096-1102.

LAMARI, N. 2009. Effect of polyunsaturated fatty acid metabolism on the eco-physiology of marine planktonic diatoms. In *Biologia Funzionale e Strutturale.* pp. 161. University of Naples Federico II, Naples.

LANG, I., HODAC, L., FRIEDL, T. & FEUSSNER, I. 2011. Fatty acid profiles and their distribution patterns in microalgae: a comprehensive analysis of more than 2000 strains from the SAG culture collection. *BMC Plant Biology* 11: 124.

LARSON, T.R. & REES, T.A.V. 1996. Changes in cell composition and lipid metabolism mediated by sodium and nitrogen availability in the marine diatom *Phaeodactylum tricornutum* (Bacillariophyceae). *Journal of Phycology* 32: 388-393.

LEFLAIVE, J. & TEN-HAGE, L. 2009. Chemical interactions in diatoms: role of polyunsaturated aldehydes and precursors. *New Phytologist* 184: 794-805.

LEWIS, J., HARRIS, A. S. D., JONES, K. J. & EDMONDS, R. L. 1999. Long-term survival of marine planktonic diatoms and dinoflagellates in stored sediment samples. *Journal of Plankton Research* 21: 343-354.

LI, W., FU, L., NIU, B., WU, S. & WOOLEY, J. 2012. Ultrafast clustering algorithms for metagenomic sequence analysis. *Briefings in Bioinformatics*: 1-13.

LOGARES, R., AUDIC, S., SANTINI, S., PERNICE, M.C., DE VARGAS, C. & MASSANA, R. 2012. Diversity patterns and activity of uncultured marine heterotrophic flagellates unveiled with pyrosequencing. *The ISME Journal* 6: 1823-1833.

LÓPEZ-GARCÍA, P. & MOREIRA, D. 2008. Tracking microbial biodiversity through molecular and genomic ecology. *Research in Microbiology* 159: 67-73.

LUND, J.W.G., MACKERETH, F.J.H. & MORTIMER, C.H. 1963. Changes in depth and time of certain chemical and physical conditions and of the standing crop of *Asterionella formosa* Hass. in the north basin of Windermere in 1947. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 246: 255-290.

LUNDHOLM, N., HASLE, G.R., FRYXELL, G.A. & HARGRAVES, P.E. 2002. Morphology, phylogeny and taxonomy of species within the *Pseudo-nitzschia americana* complex (Bacillariophyceae) with descriptions of the two new species *P. brasiliana* and *P. linea*. *Phycologia* 41: 480-497.

LUNDHOLM, N., MOESTRUP, Ø., HASLE, G.R. & HOEF-EMDEN, K. 2003. A study of the *Pseudo-nitzschia pseudodelicatissima/cuspidata* complex (Bacillariophyceae): what is *P. pseudodelicatissima*? *Journal of Phycology* 39: 797-813.

LUNDHOLM, N., MOESTRUP, Ø., KOTAKI, Y., HOEF-EMDEN, K., SCHOLIN, C. & MILLER, P. 2006. Inter- and intraspecific variation of the *Pseudo-nitzschia delicatissima* complex (Bacillariophyceae) illustrated by rRNA probes, morphological data and phylogenetic analyses. *Journal of Phycology* 42: 464-481.

LUNDHOLM, N., BATES, S.S., BAUGH, K.A., BILL, B.D., CONNELL, L.B., LÉGER, C. & TRAINER, V.L. 2012. Cryptic and pseudo-cryptic diversity in diatoms—with descriptions of *Pseudo-nitzschia hasleana* sp. nov. and *P. fryxelliana* sp. nov. *Journal of Phycology* 48: 436-454.

MACEL, M., VAN DAM, N.M. & KEURENTJES, J.J.B. 2010. Metabolomics: the chemistry between ecology and genetics. *Molecular Ecology Resources* 10: 583-593.

MANN, D. & DROOP, S. 1996. Biodiversity, biogeography and conservation of diatoms. *Hydrobiologia* 336: 19-32.

MANN, D.G. 1996a. The systematics of Stauroneis (Bacillariophyta): Sexual reproduction and auxospore development in S. anceps var. siberica *Nova Hedwigia Beiheft* 112: 307-319.

MANN, D.G. 1996b. Chloroplast morphology, movements and inheritancer in diatoms. In *Cytology, genetics and molecular biology of algae.* (Ed. by B.R. Chaudhary & S.B. Agrawal), pp. 249-274. SPB Academic Publishing bv.

MANN, D.G. 1999. The species concept in diatoms. *Phycologia* 38: 437-495.

MANN, D.G., CHEPURNOV, V.A. & DROOP, S.J.M. 1999. Sexuality, incompatibility, size variation, and preferencial polyandry in natural populations and clones of *Sellaphora pupula* (Bacillariophyceae). *Journal of Phycology* 35: 152-170.

MARGALEF, R. 1978. Life-forms of phytoplankton as survival alternatives in an unstable environment. *Oceanologica Acta* 1: 493-509.

MARGALEF, R. 1989. Size of centric diatoms as ecological indicators. *Mitteilungen der Internationale Vereinigte Limnologie* 17: 203-210.

MARTIN, J.H., COALE, K.H., JOHNSON, K.S., FITZWATER, S.E., GORDON, R.M., TANNER, S.J., HUNTER, C.N., ELROD, V.A., NOWICKI, J.L., COLEY, T.L., BARBER, R.T., LINDLEY, S., WATSON, A.J., VAN SCOY, K., LAW, C.S., LIDDICOAT, M.I., LING, R., STANTON, T., STOCKEL, J., COLLINS, C., ANDERSON, A., BIDIGARE, R., ONDRUSEK, M., LATASA, M., MILLERO, F.J., LEE, K., YAO, W., ZHANG, J.Z., FRIEDERICH, G., SAKAMOTO, C., CHAVEZ, F., BUCK, K., KOLBER, Z., GREENE, R., FALKOWSKI, P., CHISHOLM, S.W., HOGE, F., SWIFT, R., YUNGEL, J., TURNER, S., NIGHTINGALE, P., HATTON, A., LISS, P. & TINDALE, N.W. 1994. Testing the iron hypothesis in ecosystems of the equatorial Pacific Ocean. *Nature* 371: 123-129.

MARTINY, J.B.H., BOHANNAN, B.J.M., BROWN, J.H., COLWELL, R.K., FUHRMAN, J.A., GREEN, J.L., HORNER-DEVINE, M.C., KANE, M., KRUMINS, J.A., KUSKE, C.R., MORIN, P.J., NAEEM, S., OVREAS, L., REYSENBACH, A.-L., SMITH, V.H. & STALEY, J.T. 2006. Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology* 4: 102-112.

MCDONALD, S.M., SARNO, D., SCANLAN, D.J. & ZINGONE, A. 2007. Genetic diversity of eukaryotic ultraphytoplankton in the Gulf of Naples during an annual cycle. *Aquatic Microbial Ecology* 50: 75–89.

MCQUOID, M.R. & HOBSON, L.A. 1995. Importance of resting stages in diatom seasonal succession. *Journal of Phycology* 31: 44-50.

MEDLIN, L.K., KOOISTRA, W.H.C.F., GERSONDE, R. & WELLBROCK, U. 1996. Evolution of the diatoms (Bacillariophyta).II. Nuclear-encoded small-subunit rRNA sequence comparisons confirm a paraphyletic origin of the centric diatoms. *Molecular Biology and Evolution* 13: 67-75.

MEDLIN, L.K. & KOOISTRA, W.H.C.F. 2010. Methods to estimate the diversity in the marine photosynthetic protist community with ilustrations from case studies: A review. *Diversity* 2: 973-1014.

MEDLIN, L.K. 2010. Pursuit of a natural classification of diatoms: An incorrect comparison of published data. *European Journal of Phycology* 45: 155 - 166.

MENDEN-DEUER, S. & LESSARD, E.J. 2000. Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. *Limnology and Oceanography* 45: 569-579.

MEUNIER, A. 1915. *Microplankton de la mer flamande.* Memories du Mesee Royal D'Histoire Naturelle de Belgique, Bruxelles. 47 pp.

MIRALTO, A., BARONE, G., ROMANO, G., POULET, S.A., IANORA, A., RUSSO, G.L., BUTTINO, I., MAZZARELLA, G., LAABIR, M., CABRINI, M. & GIACOBBE, M.G. 1999. The insidious effect of diatoms on copepod reproduction. *Nature* 402: 173-176.

MONTAGNES, D.J.S. 1996. Growth responses of planktonic ciliates in the genera *Strobilidium* and *Strombidium*. *Marine Ecology Progress Series* 130: 241-254.

MONTAGNES, D.J.S. & FRANKLIN, D.J. 2001. Effect of temperature on diatom volume, growth rate, and carbon and nitrogen content: Reconsidering some paradigms. *Limnology and Oceanography* 46: 2008-2018.

MONTES-HUGO, M., DONEY, S.C., DUCKLOW, H.W., FRASER, W., MARTINSON, D., STAMMERJOHN, S.E. & SCHOFIELD, O. 2009. Recent changes in phytoplankton communities associated with rapid regional climate change along the Western Antarctic Peninsula. *Science* 323: 1470-1473.

MONTRESOR, M., SGROSSO, S., PROCACCINI, G. & KOOISTRA, W.H.C.F. 2003. Intraspecific diversity in *Scrippsiella trochoidea* (Dinophyceae): evidence for cryptic species. *Phycologia* 42: 56-70.

MONTRESOR, M. & LEWIS, J. 2006. Phases, stages, and shifts in the life cycles of marine phytoplankton. In *Algal cultures, analogues of blooms and applications.* (Ed. by D.V. Subba Rao), pp. 91-129. Science Publishers, Enfield.

MONTSANT, A., ALLEN, A.E., COESEL, S., MARTINO, A.D., FALCIATORE, A., MANGOGNA, M., SIAUT, M., HEIJDE, M., JABBARI, K., MAHESWARI, U., RAYKO, E., VARDI, A., APT, K.E., BERGES, J.A., CHIOVITTI, A., DAVIS, A.K., THAMATRAKOLN, K., HADI, M.Z., LANE, T.W., LIPPMEIER, J.C., MARTINEZ, D., PARKER, M.S., PAZOUR, G.J., SAITO, M.A., ROKHSAR, D.S., ARMBRUST, E.V. & BOWLER, C. 2007. Identification and comparative genomic analysis of signaling and regulatory components in the diatom Thalassiosira pseudonana1
doi:10.1111/j.1529-8817.2007.00342.x *Journal of Phycology* 43: 585-604.

MOORE, L.R., ROCAP, G. & CHISHOLM, S.W. 1998. Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. *Nature* 393: 464-467.

MORGAN-KISS, R., IVANOV, A.G., WILLIAMS, J., MOBASHSHER, K. & HUNER, N.P.A. 2002. Differential thermal effects on the energy distribution between photosystem II and photosystem I in thylakoid membranes of a psychrophilic and a mesophilic alga. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1561: 251-265.

MOUSTAFA, A., BESZTERI, B., MAIER, U.G., BOWLER, C., VALENTIN, K. & BHATTACHARYA, D. 2009. Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science* 324: 1724-1726.

MUNNS, R. & TESTER, M. 2008. Mechanisms of salinity tolerance. *Annual Review of Plant Biology* 59: 651-681.

MURRAY, J., APPELLÖF, J.J.A., GRAN, H.H., HELLAND-HANSEN, B. & HJORT, J. 1912. *The depths of the ocean: a general account of the modern science of oceanography based largely on the scientific researches of the Norwegian steamer Michael Sars in the North Atlantic / by Sir John Murray and Dr. Johan Hjort; with contributions from Professor A. Appellöf, Professor H. H. Gran and Dr. B. Helland-Hansen.* Macmillan, London. 821 pp.

NAGAI, S., HORI, Y., MANABE, T. & IMAI, I. 1995. Restoration of cell size by vegetative cell enlargement in *Coscinodiscus wailesii* (Bacillariophyceae). *Phycologia* 34: 533-535.

NAVINER, M., BERGÉ, J.P., DURAND, P. & LE BRIS, H. 1999. Antibacterial activity of the marine diatom *Skeletonema costatum* against aquacultural pathogens. *Aquaculture* 174: 15-24.

NEBEL, M., PFABEL, C., STOCK, A., DUNTHORN, M. & STOECK, T. 2011. Delimiting operational taxonomic units for assessing ciliate environmental diversity using small-subunit rRNA gene sequences. *Environmental Microbiology Reports* 3: 154-158.

NELLES, L., FANG, B.-L., VOLCKAERT, G., VANDENBERGHE, A. & WACHTER, R.D. 1984. Nucleotide sequence of a crustacean 18S ribosomal RNA gene and secondary structure of eukaryotic small subunit ribosoraal RNAs. *Nucleic Acids Research* 12: 8749-8768.

NELSON, D.M., TRÉGUER, P., BRZEZINSKI, A., LEYNAERT, A. & QUÉGUINER, B. 1995. Production and dissolution of biogenic silica in the ocean: revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biogeochemical Cycles* 9: 359-372.

NIU, B., FU, L., SUN, S. & LI, W. 2010. Artificial and natural duplicates in pyrosequencing reads of metagenomic data. *BMC Bioinformatics* 11: 187.

NODDER, S.D. & WAITE, A.M. 2001. Is Southern Ocean organic carbon and biogenic silica export enhanced by iron-stimulated increases in biological production? Sediment trap results from SOIREE. *Deep Sea Research Part II: Topical Studies in Oceanography* 48: 2681-2701.

OLSON, R.J., CHISHOLM, S.W., ZETTLER, E.R., ALTABET, M.A. & DUSENBERRY, J.A. 1990. Spatial and temporal distributions of prochlorophyte picoplankton in the North Atlantic Ocean. *Deep Sea Research Part A. Oceanographic Research Papers* 37: 1033-1051.

PARKER, M.S., MOCK, T. & ARMBRUST, E.V. 2008. Genomic insights into marine microalgae. *Annual Review of Genetics* 42: 619-645.

PIERSON, J.J., HALSBAND-LENK, C. & LEISING, A.W. 2005. Reproductive success of *Calanus pacificus* during diatom blooms in Dabob Bay, Washington. *Progress In Oceanography* 67: 314-331.

POHNERT, G. 2000. Wound-Activated Chemical Defense in Unicellular Planktonic Algae. *Angewandte Chemie International Edition* 39: 4352-4354.

PÖRTNER, H.O. & FARRELL, A.P. 2008. Physiology and Climate Change. *Science* 322: 690-692.

POSADA, D. & CRANDALL, K.A. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14: 818-818.

POSADA, D. & CRANDALL, K.A. 2001. Selecting the best-fit model of nucleotide substitution. *Systematic Biology* 50: 580-601.

PROSHKINA-LAVRENKO, A.I. 1955. *Plankton diatoms of the Black Sea.* USSR AS., Moscow. 222 pp.

PUGNETTI, A., CAMATTI, E., MANGONI, O., MORABITO, G., OGGIONI, A. & SAGGIOMO, V. 2006. Phytoplankton production in Italian freshwater and marine ecosystems: State of the art and perspectives. *Chemistry and Ecology* 22: S49-S69.

QUINCE, C., LANZEN, A., CURTIS, T.P., DAVENPORT, R.J., HALL, N., HEAD, I.M., READ, L.F. & SLOAN, W.T. 2009. Accurate determination of microbial diversity from 454 pyrosequencing data. *Nature Methods* 6: 639-641.

QUINCE, C., LANZEN, A., DAVENPORT, R. & TURNBAUGH, P. 2011. Removing Noise From Pyrosequenced Amplicons. *BMC Bioinformatics* 12: 38.

RAMBAUT, A. 1996-2002. Sequence Alignment Editor v2.0a11. In *Available at http://evolve.zoo.ox.ac.uk/ searched on 18 of January 2005.*

RAVEN, J.A. 1983. The transport and function of silicon in plants. *Biological Reviews* 58: 179-207.

RAVEN, J.A. & GEIDER, R.J. 1988. Temperature and algal growth. *New Phytologist* 110: 441-461.

RAVEN, J.A. & GEIDER, R.J. 2003. Adaptation, acclimiation and regulation in algal photosynthesis In *Photosynthesis in algae.* (Ed. by A.W. Larkum, S.E. Douglas & J.A. Raven), pp. 386–412. Kluwer academic publishers.

RAVEN, J.A. & WAITE, A.M. 2004. The evolution of silification in diatoms: inescapable sinking and sinking as escape? *New Phytologist* 162: 45-61.

REEDER, J. & KNIGHT, R. 2010. Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. *Nature Methods* 7: 668-669.

RENAUD, S.M., ZHOU, H.C., PARRY, D.L., THINH, L.-V. & WOO, K.C. 1995. Effect of temperature on the growth, total lipid content and fatty acid composition of recently isolated tropical microalgae *Isochrysis sp., Nitzschia closterium, Nitzschia paleacea,* and commercial species *Isochrysis sp.* (clone T.ISO). *Journal of Applied Phycology* 7: 595-602.

RENAUD, S.M., THINH, L.-V., LAMBRINIDIS, G. & PARRY, D.L. 2002. Effect of temperature on growth, chemical composition and fatty acid composition of tropical Australian microalgae grown in batch cultures. *Aquaculture* 211: 195-214.

RENBERG, L., JOHANSSON, A.I., SHUTOVA, T., STENLUND, H., AKSMANN, A., RAVEN, J.A., GARDESTRÖM, P., MORITZ, T. & SAMUELSSON, G. 2010. A metabolomic approach to study major metabolite changes during acclimation to limiting CO2 in *Chlamydomonas reinhardtii*. *Plant Physiology* 154: 187-196.

RENGEFORS, K. & ANDERSON, D.M. 1998. Environmental and endogenous regulation of cyst germination in two freshwater dinoflagellates. *Journal of Phycology* 34: 568-577.

REYES-PRIETO, A., HACKETT, J.D., SOARES, M.B., BONALDO, M.F. & BHATTACHARYA, D. 2006. Cyanobacterial contribution to algal nuclear genomes is primarily limited to plastid functions. *Current biology* 16: 2320-2325.

REYNOLDS, C.S. 1998. What factors influence the species composition of phytoplankton in lakes of different trophic status? *Hydrobiologia* 369/370: 11-26.

REYNOLDS, C.S. 2001. Emergence in pelagic communities. *Scientia Marina* 65.

RIBALET, F., WICHARD, T., POHNERT, G., IANORA, A., MIRALTO, A. & CASOTTI, R. 2007. Age and nutrient limitation enhance polyunsaturated aldehyde production in marine diatoms. *Phytochemistry* 68: 2059-2067.

RIBALET, F., VIDOUDEZ, C., CASSIN, D., POHNERT, G., IANORA, A., MIRALTO, A. & CASOTTI, R. 2009. High plasticity in the production of diatom-derived polyunsaturated aldehydes under nutrient limitation: physiological and ecological Iimplications. *Protist* 160: 444-451.

RIBERA D'ALCALÀ, M., CONVERSANO, F., CORATO, F., LICANDRO, P., MANGONI, O., MARINO, D., MAZZOCCHI, M.G., MODIGH, M., MONTRESOR, M., NARDELLA, M., SAGGIOMO, V., SARNO, D. & ZINGONE, A. 2004. Seasonal patterns in plankton communities in a pluriannual time series at a coastal Mediterranean site (Gulf of Naples): an attempt to discern recurrences and trends. *Scientia Marina* 68 65-83.

RILEY, G.A. & CONOVER, S.M. 1967. Phytoplankton of Long Island Sound 1954 -1955. In *Bulletin of the Bingham Oceanographic Collection.* pp. 5-34. Yale University, New York.

RINES, J.E.B. & HARGRAVES, P.E. 1990. Morphology and taxonomy of *Chaetoceros compressus* Lauder var. *hirtisetus* var. *nova*, with preliminary consideration of closely related taxa. *Diatom Research* 5: 113-127.

RIVERA, P., CRUCES, F. & CLEMENT, A. 2002. *Leptocylindrus minimus* Gran (Bacillariophyceae): morphology and distribution in Chile. *Gayana Botánica* 59: 7-11.

ROESSLER, P.G. 1988. Effects of silicon deficiency on lipid composition and metabolism in the diatom *Cyclotella cryptica*. *Journal of Phycology* 24: 394-400.

ROMANO, G., MIRALTO, A. & IANORA, A. 2010. Teratogenic effects of diatom metabolites on sea urchin *Paracentrotus lividus* embryos. *Marine Drugs* 8: 950-967.

ROONEY, A.P. 2004. Mechanisms underlying the evolution and maintenance of functionally heterogeneous 18S rRNA genes in apicomplexans. *Molecular Biology and Evolution* 21: 1704-1711.

ROSE, J.M., FENG, Y., DITULLIO, G.R., DUNBAR, R.B., HARE, C.E., LEE, P.A., LOHAN, M., LONG, M., SMITH JR., W.O., SOHST, B., TOZZI, S., ZHANG, Y. & HUTCHINS, D.A. 2009. Synergistic effects of iron and temperature on Antarctic phytoplankton and microzooplankton assemblages. *Biogeosciences* 6: 3131 - 3147.

ROUND, F., CRAWFORD, R. & MANN, D. 1990. *The Diatoms: Biology and morphology of the genera.* Cambridge University Press, Cambridge. pp.

ROUSCH, J.M., BINGHAM, S.E. & SOMMERFELD, M.R. 2003. Changes in fatty acid profiles of thermo-intolerant and thermo-tolerant marine diatoms during temperature stress. *Journal of Experimental Marine Biology and Ecology* 295: 145-156.

RUIVO, M., AMORIM, A. & CARTAXANA, P. 2011. Effects of growth phase and irradiance on phytoplankton pigment ratios: implications for chemotaxonomy in coastal waters. *Journal of Plankton Research* 33: 1012-1022.

SANCHEZ-PUERTA, M.V. & DELWICHE, C.F. 2008. A hypothesis for plastid evolution in chromalveolates. *Journal of Phycology* 44: 1097-1107.

SANTOFERRARA, L.F., MCMANUS, G.B. & ALDER, V.A. in press. Utility of genetic markers and morphology for species discrimination within the order Tintinnida (Ciliophora, Spirotrichea). *Protist*: 10.1016/j.protis.2011.12.002.

SARNO, D., KOOISTRA, W.C.H.F., MEDLIN, L.K., PERCOPO, I. & ZINGONE, A. 2005. Diversity in the genus *Skeletonema* (Bacillariophyceae). II. An assessment of the taxonomy of *S. costatum*-like species, with the description of four new species. *Journal of Phycology* 41: 151-176.

SARNO, D., KOOISTRA, W.C.H.F., BALZANO, S., HARGRAVES, P.E. & ZINGONE, A. 2007. Diversity in the genus *Skeletonema* (Bacillariophyceae): III. Phylogenetic position and morphological variability of *Skeletonema costatum* and *Skeletonema grevillei*, with the description of *Skeletonema ardens* sp. nov. *Journal of Phycology* 43: 156-170.

SARTHOU, G., TIMMERMANS, K.R., BLAIN, S. & TREGUER, P. 2005. Growth physiology and fate of diatoms in the ocean: a review. *Journal of Sea Research* 53: 25-42.

SAYEGH, F.A.Q. & MONTAGNES, D.J.S. 2011. Temperature shifts induce intraspecific variation in microalgal production and biochemical composition. *Bioresource Technology* 102: 3007-3013.

SCHANTZ, E.J., LYNCH, J.M., VAYVADA, G., MATSUMOTO, K. & RAPOPORT, H. 1966. The purification and characterization of the poison produced by *Gonyaulax catenella* in axenic culture. *Biochemistry* 5: 1191-1195.

SCHIEBER, J., KRINSLEY, D. & RICIPUTI, L. 2000. Diagenetic origin of quartz silt in mudstones and implications for silica cycling. *Nature* 406: 981-985.

SCHLOTTERER, C. 2004. The evolution of molecular markers - just a matter of fashion? *Nature Reviews Genetics* 5: 63-69.

SCHNEIDER, C., PRATT, D.A., PORTER, N.A. & BRASH, A.R. 2007. Control of oxygenation in lipoxygenase and cyclooxygenase catalysis. *Chemistry & biology* 14: 473-488.

SHAW, B.A., ANDERSEN, R.J. & HARRISON, P.J. 1997. Feeding deterrent and toxicity effects of apo-fucoxanthinoids and phycotoxins on a marine copepod *Tigriopus californicus*. *Marine Biology* 128: 273-280.

SINGER, T.L. 1998. Roles of hydrocarbons in the recognition systems of insects. *American Zoologist* 38: 394-405.

SKVORTZOW, B.W.S., B.V.). 1931. Pelagic diatoms of Korean strait of the Sea of Japan. *Philippine Journal of Science* 46: 95-122, 10 pls.

SMAYDA, T. 1980. Phytoplankton species succession In *The physiological ecology of phytoplankton.* (Ed. by I. Morris), pp. 493-570. Blackwell Scientific Publications, Oxford.

SMAYDA, T.J. 1958. Biogeographical studies of marine phytoplankton. *Oikos* 9: 158-191.

SMAYDA, T.J. 2000. Ecological features of harmful algal blooms in coastal upwelling systems. *South African Journal of Marine Science* 22: 219-253.

SOLOVCHENKO, A., KHOZIN-GOLDBERG, I., DIDI-COHEN, S., COHEN, Z. & MERZLYAK, M. 2008. Effects of light intensity and nitrogen starvation on growth, total fatty acids and arachidonic acid in the green microalga *Parietochloris incisa*. *Journal of Applied Phycology* 20: 245-251.

SOURNIA, A. 1982. Form and function in marine phytoplankton. *Biological Reviews* 57: 347-394.

STEINKE, M., STEFELS, J. & STAMHUIS, E. 2006. Dimethyl sulfide triggers search behavior in copepods. *Limnology and Oceanography* 51: 1925-1930.

STOECK, T., BEHNKE, A., CHRISTEN, R., AMARAL-ZETTLER, L., RODRIGUEZ-MORA, M., CHISTOSERDOV, A., ORSI, W. & EDGCOMB, V. 2009. Massively parallel tag sequencing reveals the complexity of anaerobic marine protistan communities. *Bmc Biology* 7: 72.

STOECK, T., BASS, D., NEBEL, M., CHRISTEN, R., JONES, M.D.M., BREINER, H.-W. & RICHARDS, T.A. 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular Ecology* 19: 21-31.

SUNDA, W., KIEBER, D.J., KIENE, R.P. & HUNTSMAN, S. 2002. An antioxidant function for DMSP and DMS in marine algae. *Nature* 418: 317-320.

SVERDRUP, H.U. 1953. On conditions for the vernal blooming of phytoplankton. *Journal du Conseil International pour l'Exploration de la Mer* 18: 287-295.

SWOFFORD, D.L. 1998. PAUP*. Phylogenetic analysis using Parimony (* and other methods). Version 4.

SYVERTSEN, E.E. 1977. *Thalassiosira rotula* and *T. gravida*, ecology and morphology. *Nova Hedwigia Beiheft* 54: 99-112.

TELFORD, R.J., VANDVIK, V. & BIRKS, H.J.B. 2006. Dispersal limitations matter for microbial morphospecies. *Science* 312: 1015.

TEOH, M.-L., CHU, W.-L., MARCHANT, H. & PHANG, S.-M. 2004. Influence of culture temperature on the growth, biochemical composition and fatty acid profiles of six Antarctic microalgae. *Journal of Applied Phycology* 16: 421-430.

THOMPSON, P.A., GUO, M.-X. & HARRISON, P.J. 1992a. Effects of variation in temperature. I. on the biochemical composition of eight species of marine phytoplankton. *Journal of Phycology* 28: 481-488.

THOMPSON, P.A., GUO, M.-X., HARRISON, P.J. & WHYTE, J.N.C. 1992b. Effects of variation in temperature. II. On the fatty acid composition of eight species of marine phytoplankton. *Journal of Phycology* 28: 488-497.

TRÉGUER, P., NELSON, D.M., VAN BENNEKOM, A.J., DEMASTER, D.J., LEYNAERT, A. & QUÉGUINER, B. 1995. The silica balance in the world ocean: a reestimate. *Science* 268: 375-379.

VAN'T HOFF, M.J.H. 1884. Etudes de dynamique chimique. *Recueil des Travaux Chimiques des Pays-Bas* 3: 333-336.

VERITY, P.G. 1981a. Effects of temperature, irradiance, and daylength on the marine diatom *Leptocylindrus danicus* Cleve. II. Excretion. *Journal of Experimental Marine Biology and Ecology* 55: 159-169.

VERITY, P.G. 1981b. Effects of temperature, irradiance, and daylength on the marine diatom *Leptocylindrus danicus* Cleve. I. Photosynthesis and cellular composition. *Journal of Experimental Marine Biology and Ecology* 55: 79-91.

VERITY, P.G. 1982a. Effects of temperature, irradiance, and daylength on the marine diatom *Leptocylindrus danicus* Cleve. IV. Growth. *Journal of Experimental Marine Biology and Ecology* 60: 209-222.

VERITY, P.G. 1982b. Effects of temperature, irradiance, and daylength on the marine diatom *Leptocylindrus danicus* Cleve. III. Dark respiration. *Journal of Experimental Marine Biology and Ecology* 60: 197-207.

VIDOUDEZ, C. & POHNERT, G. 2008. Growth phase-specific release of polyunsaturated aldehydes by the diatom *Skeletonema marinoi*. *Journal of Plankton Research* 30: 1305-1313.

VIDOUDEZ, C. & POHNERT, G. 2011. Comparative metabolomics of the diatom *Skeletonema marinoi* in different growth phases. *Metabolomics*: 1-16.

VOLKMAN, J.K., JEFFREY, S.W., NICHOLS, P.D., ROGERS, G.I. & GARLAND, C.D. 1989. Fatty acid and lipid composition of 10 species of microalgae used in mariculture. *Journal of Experimental Marine Biology and Ecology* 128: 219-240.

VON STOSCH, H.A. 1965. Manipulierung der zellgrösse von diatomeen in Experiment. *Phycologia* 5: 21-44.

WERNER, D. & STANGIER, E. 1976. Silica and temperature dependent colony size of *Bellerochea maleus* f. *biangulata* (Centrales, Diatomeae). *Phycologia* 15: 73-77.

WHITAKER, R.J., GROGAN, D.W. & TAYLOR, J.W. 2003. Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* 301: 976-978.

WICHARD, T., POULET, S., HALSBAND-LENK, C., ALBAINA, A., HARRIS, R.P., LIU, D. & POHNERT, G. 2005a. Survey of the chemical defence potential of diatoms: screening of fifty one species for a, b, c, d-unsaturated aldehydes. *Journal of Chemical Ecology* 31: 949-958.

WICHARD, T., POULET, S.A., HALSBAND-LENK, C., ALBAINA, A., HARRIS, R., LIU, D. & POHNERT, G. 2005b. Survey of the chemical defence potential of diatoms: screening of fifty species for α,β,γ,δ-unsaturated aldehydes. *Journal of Chemical Ecology* 31: 949-958.

WOESE, C.R. & FOX, G.E. 1977. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences* 74: 5088-5090.

WOLFE-SIMON, F., STAROVOYTOV, V., REINFELDER, J.R., SCHOFIELD, O. & FALKOWSKI, P.G. 2006. Localization and role of manganese superoxide dismutase in a marine diatom. *Plant Physiology* 142: 1701-1709.

WU, D., HARTMAN, A., WARD, N. & EISEN, J.A. 2008. An automated phylogenetic tree-based small subunit rRNA taxonomy and alignment pipeline (STAP). *PLoS One* 3: e2566.

YODER, J.A. 1979. Effect of temperature on light/limited growth and chemical composition of *Skeletonema costatum* (Bacillariophyceae). *Journal of Phycology* 15: 362-370.

YOON, H.S., HACKETT, J.D., CINIGLIA, C., PINTO, G. & BHATTACHARYA, D. 2004. A molecular timeline for the origin of photosynthetic eukaryotes. *Molecular Biology and Evolution* 21: 809-818.

YU, E., ZENDEJAS, F., LANE, P., GAUCHER, S., SIMMONS, B. & LANE, T. 2009. Triacylglycerol accumulation and profiling in the model diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum* (Baccilariophyceae) during starvation. *Journal of Applied Phycology* 21: 669-681.

ZAPATA, M., RODRÍGUEZ, F., FRAGA, S., BARRA, L. & RUGGIERO, M.V. 2011. Chlorophyll *c* pigment patterns in 18 species (51 strains) of the Genus *Psedo-nitzschia* (Bacillariophyceae). *Journal of Phycology* 47: 1274-1280.

ZIELINSKI, U. & GERSONDE, R. 1997. Diatom distribution in Southern Ocean surface sediments (Atlantic sector): Implications for paleoenvironmental reconstructions. *Palaeogeography, Palaeoclimatology, Palaeoecology* 129: 213-250.

ZINGONE, A., MONTRESOR, M. & MARINO, D. 1990. Summer phytoplankton physiognomy in coastal water of the Gulf of Naples. *P.S.Z.N.I: Marine Ecology* 11: 157-172.

ZINGONE, A., CASOTTI, R., RIBERA D'ALCALÀ, M., SCARDI, M. & MARINO, D. 1995. 'St Martin's Summer': the case of an autumn phytoplankton bloom in the Gulf of Naples (Mediterranean Sea) *Journal of Plankton Research* 17: 575-593.

ZINGONE, A., PERCOPO, I., SIMS, P.A. & SARNO, D. 2005. Diversity in the genus *Skeletonema* (Bacillariophyceae). I. A re-examination of the type material of *Skeletonema costatum*, with the description of *S. grevillei* sp. nov. *Journal of Phycology* 41: 140-150.

ZINGONE, A. & WYATT, T. 2005. Harmful algal blooms: keys to the understanding of the phytoplankton ecology. In *The Sea.* (Ed. by A.R. Robinson & K.H. Brink), pp. 867-926. Harvard University Press, Harvard.

Before you start some work, always ask yourself three questions - Why am I doing it, What the results might be and Will I be successful. Only when you think deeply and find satisfactory answers to these questions, go ahead - Chanakya

Legatoria - Tipografia RICCO - Via Mezzocannone, 105 Napoli
Tel./Fax 081.551.66.55 e-mail: tipografiaricco@libero.it
Premiata con Medaglia d'oro "Premio Minerva 1988"